DIPLOMARBEIT

Titel der Diplomarbeit

Differences within the superfamily of Chlorite Dismutases: structural and functional characterization

Verfasser

Georg Mlynek

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag.rer.nat.)

Wien, 2010

meiner Familie

Contents

Chapter I

General Introduction and Outline
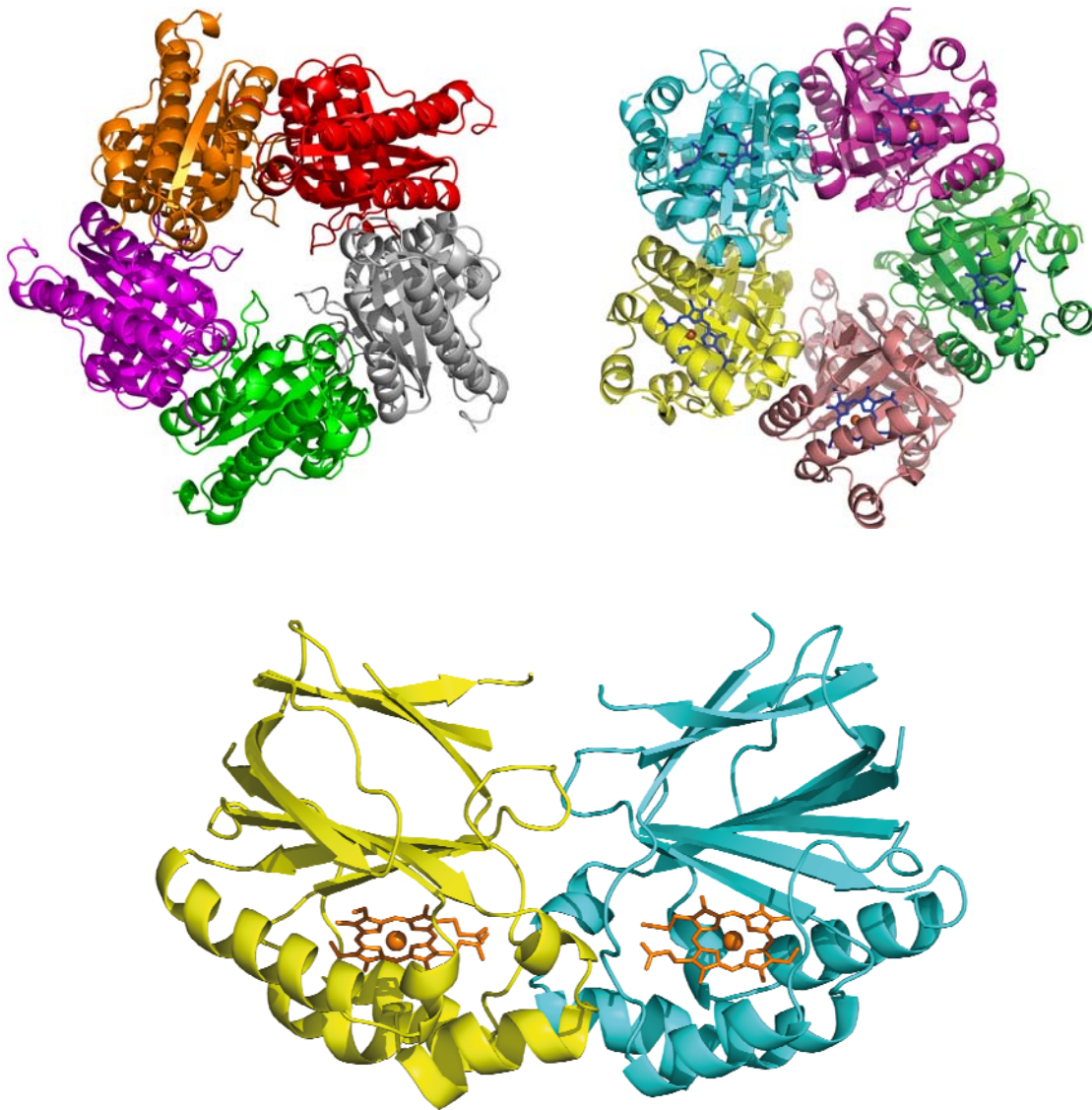
Front: Cartoon representation of chlorite dismutases from *Listeria monocytogenes (LmCld)*, *"Candidatus Nitrospira defluvii" (NdCld)* and *Nitrobacter winogradsky* (NwCld) (from left to right). The structures were solved and analysed during this thesis and brought new insight in the superfamily of chlorite dismutases.

General Introduction

Oxochlorates

Perchlorate (ClO_4^-) and its reduced oxy-anionic derivatives contaminats of soil-, surface- and ground-waters are becoming a serious threat to human health and to the environment.

Except for the relatively high natural abundances of perchlorate in Chilean caliche (precipitated salts in soil from evaporated wetting fronts) and in some potash ores, naturally deposits tend to be of low perchlorate concentrations. Recent studies show that perchlorate can be formed in the atmosphere, which would explain the high distribution of perchlorate reducing bacteria species (Furdui and Tomassini ; Dasgupta, Martinelango et al. 2005). Manmade deposits – typically sites for manufacturing of perchlorate compounds and manufacturing, testing, or disposal of solid rocket propellant usually contain high concentration of perchlorate.

Tonacchera et al. (Tonacchera, Pinchera et al. 2004) could show that the human sodium/iodide symporter (NIS) has a 30-fold higher affinity for perchlorate than for iodide. As iodide is needed for biosynthesis of the iodine-containing hormones thyroxine and triiodothyronine, which are essential for growth, development, and many metabolic reactions, perchlorate posses a great threat to human health (Tran, Valentin-Blasini et al. 2008).

Chlorate (ClO_3^-) is used as a weed controller, defoliant and soil steriliant in agriculture as well as for on-site generation of chlorine dioxide (ClO_2) which is used as a bleaching agent in the pulp and paper industry and in oxidative treatment of drinking water.

Chlorite (ClO_2^-) is mainly used for producing chlorine dioxide.

Perchlorate, chlorate and chlorite are very persistent and their removal with chemical and physical water treatment technologies is still inefficient. Thus bioremediation is the method of choice (Wu, He et al. 2008): several microorganisms can reduce oxyanions of chlorite such as chlorate (ClO_3^-) and perchlorate (ClO_4^-) under anaerobic conditions.

The total reduction of perchlorate requires two enzymes: perchlorate reductase and chlorite dismustase. Perchlorate reductase is a molybdopterin-dependent enzyme belonging to the DMSO reductase family of enzymes. It catalyzes the first two steps of reduction process, the conversion of ClO_4^- to ClO_2^- (Fig. 1). Chlorite dismutase (Cld) catalyzes the third step, the reduction of toxic chlorite to innocuous Cl^- and O_2 (Ueno, Oishi et al. 2000) (Fig. 1).
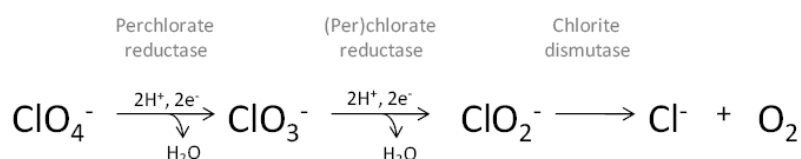
Figure 1 **(Per)chlorate reduction pathway**

Chlorite dismutase

As proposed by Hagendoorn et al. the correct name of this enzyme should be chloride:oxygen oxidoreductase or chlorite O2-lyase (EC1.13.11.49), because the reaction is not a dismutation or disproportionation of chlorite, but an intramolecular redox reaction.

Lee *et al.* (Lee, Streit et al. 2008) proposed a reaction mechanism where compound I (an oxoiron(IV) porphyrin π-cation radical) and hypochlorite (ClO^-) are formed in the first reaction step during the reaction of chlorite with the ferric enzyme. Subsequently, hypochlorit makes a nucleophilic attack on to compound I and compound II is generated which decays to chloride and oxygen (Fig. 2).

Figure 2 **Reaction pathway of chlorite degradation by Cld as proposed by Lee et al**.

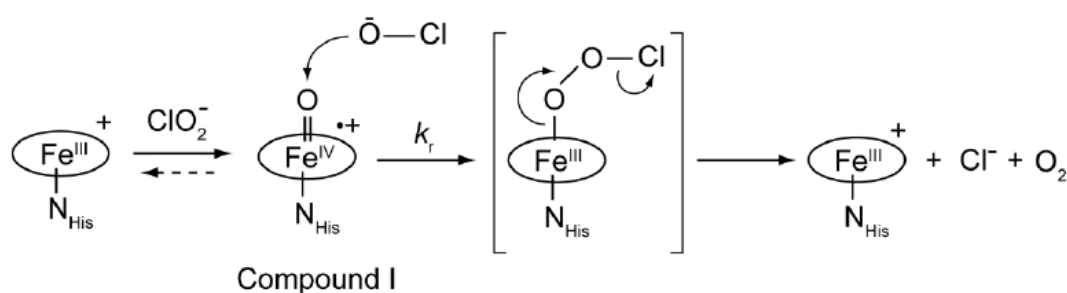The heme containing enzyme chlorite dismutase was first described by van Ginkel, Rikken et al in 1996 (van Ginkel, Rikken et al. 1996). The enzyme was discovered in strain GR-1 (now *Azospira oryzae* GR-1) which is a perchlorate-reducing bacterium (PCB) and uses ClO_3^- or ClO_4^- as alternative electron acceptors in absence of oxygen (van Ginkel, Rikken et al. 1996)

Frank Maixner et al. recently found a highly active Cld in the nitrite-oxidizing bacterium "*Candidatus* Nitrospira defluvii" (Maixner, Wagner et al. 2008) and showed that Clds and related proteins are widely distributed among the *Bacteria* and *Archaea*. This is very surprising because the presence of Clds was until recently restricted to organisms capable of the energy-gaining respiration of (per)chlorate .

Aim of this thesis

The aim of this thesis was to gain new insights in the recently discovered superfamily of chlorite dismutases using a comprehensive structural, bioinformatical and biochemical approach.

For this reason we choose the Cld from "*Candidatus* Nitrospira defluvii" (NdCld) a key-nitrifier in biological wastewater treatment and key-player in the global nitrogen cycle because no Cld from a non-proteobacterial organism was analyzed before. The study on NdCld focused mainly on the reaction mechanism.

The Cld of Nitrobacter winogradsky (NwCld) which sequence was found to be shorter than sequences of previously analysed Cld and which like"*Candidatus* Nitrospira defluvii" is also a key-player in wastewater treatment plants and the global nitrogen cycle is a superb candidate for experimentally evaluating whether the phylogenetic and structural diversity of efficient Clds might be greater than previously anticipated and should strengthen hypothesis we made on the reaction mechanism using NdCld in the second chapter.

To gain more insights into the lineage of Cld-like enzymes, we took the Cld from Listeria monocytogenes (LmCld) that was found to be essential for this facultative intracellular pathogenic bacteria. As there are no Cld-homologs in humans which could cause adverse reactions, LmCld is a promising drug target against this serious human pathogen.

A summary of the presented studies is given in **Chapter V**.

Chapter II

Structural and functional characterisation of the chlorite dismutase from the nitriteoxidizing bacterium *"Candidatus* Nitrospira defluvii"*: Identification of a catalytically important amino acid residue

Accepted Manuscript

Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium " *Candidatus* Nitrospira defluvii": Identification of a catalytically important amino acid residue

Julius Kostan, Björn Sjöblom, Frank Maixner, Georg Mlynek, Paul Georg Furtmüller, Christian Obinger, Michael Wagner, Holger Daims, Kristina Djinović-Carugo

Please cite this article as: Kostan, J., Sjöblom, B., Maixner, F., Mlynek, G., Furtmüller, P.G., Obinger, C., Wagner, M., Daims, H., Djinović-Carugo, K., Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium " *Candidatus* Nitrospira defluvii": Identification of a catalytically important amino acid residue, *Journal of Structural Biology* (2010), doi: 10.1016/j.jsb.2010.06.014

Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium "*Candidatus* Nitrospira defluvii": Identification of a catalytically important amino acid residue

Julius Kostan[1], Björn Sjöblom[1], Frank Maixner[2,5], Georg Mlynek[1], Paul Georg Furtmüller[3], Christian Obinger[3], Michael Wagner[2], Holger Daims[2,*] and Kristina Djinović-Carugo[1,4,*]

[1]Department for Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, A-1030 Vienna, Austria; [2]Department of Microbial Ecology, Vienna Ecology Centre, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria; [3]Department of Chemistry, Division of Biochemistry, BOKU – University of Natural Resources and Applied Life Sciences, Muthgasse 18, Vienna, A-1190, Austria; [4]Department of Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, 1000 Ljubljana, Slovenia; [5]present address: Institute for Mummies and the Iceman, EURAC research, Viale Druso 1, 39100 Bolzano, Italy

Running title: Structure of chlorite dismutase from *Nitrospira defluvii*
***Address correspondence to**: Kristina Djinović-Carugo, Email: kristina.djinovic@univie.ac.at, Tel.: +43-1-4277-52203; Fax: +43-1-4277-9522; Holger Daims, Email: daims@microbial-ecology.net, Tel.: +43-1-4277-54392; Fax: +43-1-4277-54389

ABSTRACT

Chlorite dismutase (Cld) is a unique heme enzyme which transforms chlorite to chloride and molecular oxygen (reaction: $ClO_2^- \rightarrow Cl^- + O_2$). Since bacteria with Cld play significant roles in the bioremediation of industrially contaminated sites and also in wastewater treatment, it is of high interest to understand the molecular mechanism of chlorite detoxification. Here we investigate a highly active Cld from *Candidatus* Nitrospira defluvii, a key nitrifier in biological wastewater treatment, using a comprehensive structural, biochemical and bioinformatics approach. We determined the crystal structure of Cld from *Candidatus* Nitrospira defluvii (NdCld) and showed that functional NdCld is a homopentamer possessing a fold found in other Clds and Cld-like enzymes. To investigate the Cld function in more detail, site-directed mutagenesis of a catalytically important residue (Arg173) was performed and two enzyme mutants were structurally and biochemically characterized. Arginine 173 is demonstrated to play a key role in (i) controlling of ligand and substrate access and binding and (ii) in chlorite dismutation reaction. The flexible residue modulates the electrostatic potential and size of the active site entrance and might be involved in keeping transiently formed hypochlorite in place for final molecular oxygen and chloride formation. Furthermore, using a structure-based sequence alignment, we show that the residue corresponding to Arg173 is conserved in all known active forms of Cld and propose it as a marker for Cld activity in yet uncharacterized Cld-like proteins. Finally, our analysis indicates that all Clds and Cld-like enzymes employ a non-covalently bound heme as a cofactor.

Keywords: Chlorite dismutase, *Candidatus* Nitrospira defluvii, 3D-structure, enzyme kinetics, structure-based alignment, site-directed mutagenesis, catalytic mechanism, heme proteins, (per)chlorate bioremediation

1

INTRODUCTION

The heme enzyme chlorite dismutase (Cld)[$] was discovered in chlorate- and perchlorate-reducing bacteria (PCRB) (van Ginkel et al., 1996), which are facultative anaerobes using ClO_3^- or ClO_4^- as alternative electron acceptors in absence of oxygen. The reduction of (per)chlorate yields chlorite (ClO_2^-), which is a strong oxidant and has cell-damaging effects (Ueno et al., 2000). At this point Cld comes into play: this enzyme degrades chlorite to Cl^- and O_2 and thus protects PCRB from the accumulation of harmful chlorite. The reaction catalyzed by Cld is not energy-releasing, but is regarded as essential for the survival of PCRB in presence of (per)chlorate. Surprisingly, a highly active Cld was recently identified, characterized, and found to be expressed *in vivo* in the nitrite-oxidizing bacterium "Candidatus Nitrospira defluvii"(Maixner et al., 2008). Unlike PCRB, nitrite-oxidizing bacteria (NOB) are aerobic and chemolithoautotrophic organisms that gain energy from the oxidation of nitrite to nitrate. To date, no NOB have been shown to grow by oxidizing nitrite with (per)chlorate as electron acceptor, although this would be energetically possible (Maixner et al., 2008). Furthermore, *Ca*. N. defluvii is a member of the bacterial phylum *Nitrospirae* (Spieck et al., 2006), whereas almost all known PCRB belong to the *Proteobacteria*. The discovery of an active Cld in *Ca*. N. defluvii (NdCld) raised fundamental questions regarding the evolution of Cld and the biological role of the enzyme in this organism indicating that NOB may possess yet unknown, and fully unexpected, ecophysiological features. Moreover, in the same study (Maixner et al., 2008) all available microbial genomes were screened for genes encoding proteins with homology to known Cld. This search revealed a relatively large superfamily of Cld-like proteins, which extends over a surprising number of bacterial and also archaeal phyla. The known *bona fide* Clds from PCRB and *Ca*. N. defluvii constitute a minor group within this superfamily, whereas the enzymatic activities of the vast majority of the Cld-like proteins are unknown. Merely one Cld-like protein from *Thermus thermophilus* has been characterized both structurally and biochemically, and it was found to have only a weak Cld activity (Ebihara et al., 2005). Cld is of high practical relevance for the bioremediation of the anthropogenic pollutants (per)chlorate and chlorite. These toxic compounds are released into the environment due to industrial processes using (per)chlorate and the application of fertilizers, disinfectants and bleaching agents containing (per)chlorate (Coates and Achenbach, 2004). In addition, Cld is particularly interesting from the biochemical perspective: besides photosystem II, Cld is the only known enzyme whose main function is to catalyze the formation of a covalent O-O bond, which leads to the synthesis of molecular oxygen. However, the catalytic mechanism of chlorite degradation by Cld is not completely understood yet. Significant progress was made recently by Lee et al. (2008), who proposed a mechanism for O-O bond formation by Cld that involves the transfer of one oxygen atom from chlorite to the heme iron, resulting in a compound I intermediate, and the subsequent recombination of hypochlorite and compound I (Eq. 1).

2

Compound I

$$\tag{1}$$
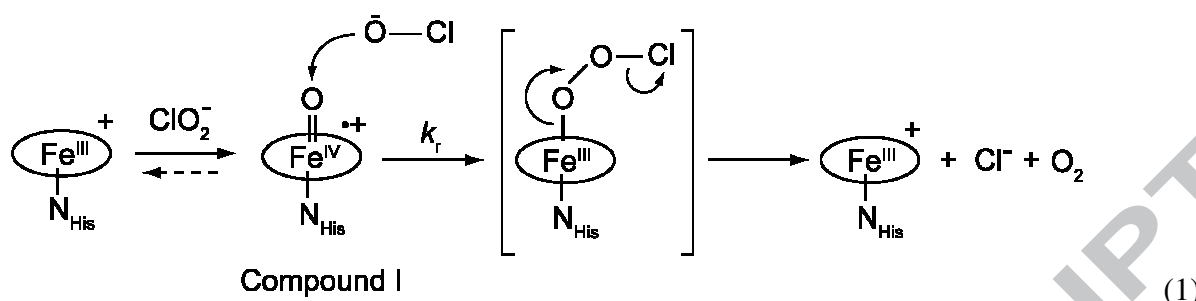
Further, valuable insight was achieved by analysis of the crystal structure of a highly active Cld, from the betaproteobacterial perchlorate reducer *Azospira oryzae* strain GR-1 (AoCld) (de Geus et al., 2009) and recently by a structure of Cld from *Dechloromonas aromatica* (DaCld) (Goblirsch et al., 2010). Here we report on the structural, bioinformatics and biochemical analyses of the Cld of the nitrite oxidizer *Ca*. N. defluvii. NdCld is the first analyzed, highly active Cld from a non-proteobacterial organism. Based on structural data and theoretical considerations, a particular amino acid residue was previously suspected to play a key role for chlorite degradation (de Geus et al., 2009; Goblirsch et al., 2010), but experimental support for its function was lacking. We addressed this question by studying site-directed mutants of NdCld at the levels of protein structure and enzyme activity. Our results confirm the functional importance of this residue and thus extend current evidence-based knowledge on the catalytic mechanism of Cld. Moreover, we show that this specific residue is conserved in all Clds known to efficiently degrade chlorite and may allow one to predict *in silico* whether yet uncharacterized Cld-like proteins possess a high Cld activity.

MATERIALS AND METHODS

Cloning

The partial gene of NdCld (without the N-terminal signal peptide) was amplified from an enrichment of *Ca*. Nitrospira defluvii (Spieck et al., 2006) by PCR and cloned into expression vector pETM11 (EMBL) for the production of a N-terminal His-tagged fusion protein as previously described (Maixner et al., 2008).

Site-directed mutagenesis

To obtain plasmids for expressing the mutated proteins NdCld R173A or NdCld R173K, where either alanine or lysine was substituted for residue Arg173, mutagenesis was carried out using the QuickChange site-directed mutagenesis kit (Stratagene) with the following primers and their reverse complements: 5´-TATCTGAAGACGGTGAAAGCAAAACTGTATCATTCGAC G-3´, and 5´-CTGAAGACGGTGAAAAAGAAACTGTATCATTCG-3´, respectively. The plasmid encoding the N-terminal His-tag fusion wild-type NdCld (without the N-terminal signal peptide) was used as template.

3

Expression and purification

Native and mutated NdClds were expressed in *E. coli* BL21(DE3) RIL cells (Stratagene) cultivated in heme-enriched (25-50 μM heme *b*) Terrific Broth medium. The N-terminal His-tagged fusion proteins were purified by affinity chromatography using HisTrap HP columns (GE Healthcare). Subsequently, the His-tag was cleaved overnight with TEV protease added at a ratio of 1:100 (w/w), and the NdCld was further purified by size exclusion chromatography (SEC) on a HiLoad 26/60 Superdex 200 pg column (GE Healthcare) equilibrated with 20 mM Hepes, 100 mM NaCl, 0.2% glycerol, pH 7.5. In order to obtain protein samples fully loaded with the heme cofactor, purified proteins were mixed with hemin (Fluka) at a molar ratio of 1:2 and incubated for 1 h at room temperature. The last purification step (size exclusion chromatography) served to remove unbound hemin from the mixture. Purified NdClds were concentrated to 15-30 mg/ml and stored at -80°C until further use.

Crystallization

Initial attempts to crystallize NdCld, which had been prepared without addition of hemin were performed at 22°C using the sitting-drop vapor diffusion technique and a nanodrop-dispensing robot (Phoenix™ RE). After 2-3 days 80% of the conditions of PACT Suite crystallization screen (Qiagen) displayed hits. Nevertheless, despite many optimization trials of several selected leads, none of them diffracted to better than 4.5 Å resolution. Subsequent crystallization screens with NdClds, which had been fully loaded with heme *b* were performed using the same methodology. The hanging-drop vapor diffusion method was used for the refinement of crystallization conditions of all forms of NdCld using 24-well Linbro plates sealed with siliconized cover slides. Drops (4 μl) containing equivalent amounts of protein and precipitant solutions were mixed and equilibrated against 0.45 ml of reservoir at 22°C. Wild-type NdCld in complex with imidazole was crystallized in 0.1 M Hepes pH 7.5, 1.4 M sodium/potassium phosphate. Imidazole was found to bind to heme during affinity purification on a HisTrap HP column, and to co-crystallize with NdCld. To obtain cyanide-bound NdCld, the crystals grown under the aforementioned conditions were soaked in 5 mM KCN for 1 hour at 22°C. All soaking steps were performed immediately before flash freezing of crystals in liquid nitrogen. The crystals of the NdCld R173A and R173K mutants were grown from 0.1 M sodium acetate pH 4.6, 1.4 M ammonium phosphate, and 0.1 M citric acid, pH 4.0, 0.4-0.6 M ammonium sulfate, respectively.

Data collection and processing

Diffraction data from NdCld crystals were collected either in-house on a Bruker Microstar rotating anode at 1.54 Å wavelength X-ray radiation or at various beamlines at ESRF. Diffration data collected in-house was integrated and scaled using the Proteum2 software suite (Bruker AXS Inc.), while the synchrotron data sets were processed using XDS (Kabsch, 1993). Data collection statistics are summarized in Table 1.

4

Phasing

Experimental phases were obtained from a data set (data set 1) collected in-house with 1.54 Å wavelength X-ray radiation using SAD-methodology, exploiting the anomalous signal of iron and sulfur atoms. Heavy atom search, SAD phasing, solvent flattening and auto-building were carried out in autoSHARP (Vonrhein et al., 2007) using SHELXD (Uson and Sheldrick, 1999), SHARP (Vonrhein et al., 2007), SOLOMON (Abrahams et al., 1996) and Arp/wARP (Morris et al., 2003) for the successive steps. All subsequent data sets were phased by molecular replacement using MOLREP (Vagin and Teplyakov, 1997) with the refined NdCld model resulting from data set 1 as a search model.

Building, refinement and validation of the structure

The structures of native NdCld as well as of the R173A or R173K mutants were refined with Refmac5 (Murshudov et al., 1997). Manual model building was performed in COOT (Emsley and Cowtan, 2004) and final validation of the models was performed with MOLPROBITY (Davis et al., 2004). Data collection and refinement statistics are summarized in Table 1.

Steady-state kinetics

The stoichiometry of the Cld reaction has been reported to be 1 mol Cl^- and 1 mol O_2 out of 1 mol ClO_2^- (Lee et al., 2008). In this work the release of molecular oxygen was continuously monitored by using a Clark-type electrode (YSI 5331 Oxygen Probe) inserted into a stirred water bath (YSI 5301B) kept at 30°C. Measurements were performed in 50 mM phosphate buffer, pH 7.0, and the electrode was equilibrated to 100% air saturation (i.e. 245 µM O_2) by bubbling air to the reaction mixture for at least 15 minutes. After removal of molecular oxygen from the buffer by bubbling with nitrogen, chlorite was added at different initial concentrations and the reaction was started by addition of 20 nM and 200 nM solutions of wild-type and mutant NdCld (R173A or R173K), respectively. Molecular oxygen production rates (µM O_2 s^{-1}) were calculated from the initial linear time traces and plotted against chlorite concentrations.

Transient-state kinetics

For these experiments we used a stopped-flow apparatus (model SX-18MV, Applied Photophysics) equipped for both conventional and sequential measurements. The optical quartz cell with a pathlength of 10 mm had a volume of 20 µl. The fastest time for mixing two solutions and recording the first data point was 1.3 ms. All measurements were performed at 25°C. Cyanide binding to ferric NdCld was measured in the conventional stopped-flow mode by following the decrease of absorbance at the Soret maximum (410 nm). In a typical experiment, one syringe contained 4 µM NdCld in 50 mM phosphate buffer, pH 7.0, and the second contained at least a 6-fold excess of cyanide in the same buffer. At least six measurements were performed for each ligand concentration. The mean of the pseudo-first-order rate constants, k_{obs}, was used in the calculation of the second-order rate constants obtained from the slope of a plot of k_{obs} *versus* cyanide concentration. In addition cyanide binding to the three heme proteins was also investigated using the diode array detector (Applied Photophysics PD.1) attached to the

5

stopped-flow machine. Normal data sets were analyzed using the Pro-K simulation program (Applied Photophysics), which allowed the synthesis of artificial sets of time-dependent spectra as well as spectral analysis of enzyme intermediates.

Structural analysis and superposition

Structure comparisons and superpositions were performed by using the SSM server (Protein Structure Comparison service SSM at EBI) and the program SUPERPOSE of the CCP4-package (CCP4, 1994; Krissinel and Henrick, 2004).

Phylogenetic analyses of Cld-like proteins

An already existing dataset of aligned amino acid sequences of 174 Clds and Cld-like proteins (Maixner et al., 2008) was manually re-aligned based on the available crystal structures of the following proteins: NdCld from *Ca*. Nitrospira defluvii; Cld-like protein from *Geobacillus stearothermophilus* (PDB: 1T0T); Cld-like protein from *Thermoplasma acidophylum* (PDB: 3DTZ); Cld-like protein from *Thermus thermophilus* (PDB: 1VDH); Cld from *Azospira oryzae* (PDB: 2VXH). Highly conserved secondary and tertiary structure motifs were used to identify homologous residues in those sequences, which are relatively dissimilar at the primary structure level. Based on the refined alignment and on the 3D structures, all amino acid residues close to the active site (less than about 4 Å from the heme) were identified, and the conservation of these residues within and between the different phylogenetic lineages of Cld-like proteins was analyzed. Phylogenetic analysis was performed with a subset of 76 of these Cld-like proteins as already described in Maixner et al. (2008). In total 243 alignment columns were used for phylogenetic analysis.

Miscellaneous

Protein concentrations were determined by using the Lowry assay with bovine serum albumin as standard. Heme type and NdCld:heme *b* stoichiometry were determined by the pyridine hemochrome assay (Berry and Trumpower, 1987), and by using the QuantiChrom™ heme assay kit (BioAssay Systems), respectively.

Protein Data Bank Accession Codes

The atomic coordinates and structure factors (codes *XXX, XXX, XXX* and *XXX*) have been deposited in the Protein data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/)

RESULTS

Overall structure

Initial experiments with full length wild-type NdCld had shown that heterologous expression of the complete *cld* gene (including the signal peptide) did not yield a functional enzyme

6

(Maixner et al., 2008). Therefore, we decided to crystallize a version of the enzyme lacking the N-terminal, 26 amino acids long signal peptide necessary for protein export into the periplasmic space. The structure of NdCld in complex with heme *b* was determined using phases from the SAD experiment exploiting Fe and S anomalous signals. Data were collected and crystal structures refined for NdCld in complex with cyanide or imidazole, as well as for two mutant forms of the enzyme, where arginine at position 173 was substituted by either alanine (R173A), or lysine residue (R173K) (for details see Table 1). While wild-type NdCld crystallized in the hexagonal space group P3$_2$21, both mutants crystallized in the monoclinic space group C2 (Table 1). The overall quaternary structure of NdCld was found to be essentially identical through all crystallized proteins, exhibiting five monomers being arranged in a ring-like fashion around a central channel (Fig. 1). This is in good agreement with the result of size exclusion chromatography, giving an approximate molecular weight of 130 kDa for wild-type enzyme and the NdCld mutants. The "donut-shaped" pentamer has an outer diameter of about 80 Å and a height of 70 Å. The central channel is approximately 20 Å in diameter. Thus, the overall structure resembles the architecture previously unveiled for the Cld-like proteins of *Thermoplasma acidophilum*, *Geobacillus stearothermophilus*, *Thermus thermophilus* HB8 and *Dechloromonas aromatica* (PDB: 3DTZ, 1T0T, 1VDH; access to DaCld 3M2S, 3M2Q is not available yet). The only exception is the structure of the Cld of *Azospira oryzae* strain GR-1 (PDB: 2VXH), for which a hexameric oligomeric state in the crystal has been reported (de Geus et al., 2009).

Subunit structure

Each monomer of NdCld is characterized by two topologically equivalent four-stranded antiparallel β-sheets forming a β-barrel, flanked on both sides by 6 α-helices (Fig. 2A). The order of strands in the β-sheet is 4↑1↓3↑2↓, resembling the topology known as the ferredoxin-like fold, which is seen also in other Cld-like enzymes (de Geus et al., 2009; Ebihara et al., 2005). Two longer helices (α3, α4) span the length of the β-sheet and run approximately parallel to the strands forming a cavity that is closed on one side by two short helices (α2, α5) which are arranged perpendicularly to the β-sheet. The two β-sheets pack together at the angle of about 65 degrees, forming a central flattened β-barrel surrounded on both sides by α-helices. Thus, each NdCld subunit is formed by two structurally similar domains that superimpose with a root mean square deviation (r.m.s.d.) of 2.1 Å over 97 Cα atoms (Fig. 2B). The main differences between the two domains are in the angles that helices α3' and α4' form with the β-sheet, resulting in a larger cavity of the C-terminal domain. In addition, in the C-terminal domain the last β-strand β4' is divided into two parts by a loop (Fig. 2A and B). Comparison of the individual subunits of the pentamer showed that the molecules are almost identical (r.m.s.d. of ~ 0.4 Å over all Cα atoms). When compared with corresponding subunits of other Cld-like enzymes (PDB: 1T0T, 3DTZ, 1VDH, 2VXH), all subunits superimposed together with a 1.72 Å r.m.s.d. over 180 pairs of structurally equivalent Cα atoms, revealing a high structural conservation of the subunit fold. NdCld showed the highest structural similarity to the AoCld (r.m.s.d. of 0.98 Å over 180 Cα atoms), which had also been crystallized in complex with heme *b* (de Geus et al., 2008). Although the structural comparison with DaCld (Goblirsch et al., 2010) was not possible, we on the basis of similarity with AoCld and NdCld believe that the overall structures will be very similar. The main difference between the structures of Clds with (AoCld, NdCld) and without

7

heme *b* (PDB: 1T0T, 3DTZ, 1VDH), is the orientation of the loop region between β4 and α1' that is partly involved in binding of the prosthetic group and thereby being stabilized in different conformation in apo and holo forms.

The common interface between neighboring subunits is 1400 Å^2. Since each subunit is in contact with two other molecules, the total area buried in the interfaces is about 23 % of the total surface area of a subunit. Each half of a subunit was found to interact with the corresponding half of a neighboring molecule (Fig. 1). The interface consists mainly of residues from helix α4 and strand β4, which interact with residues in the loop between strands β2 and β3 of the neighboring subunit. The compact organization of the NdCld pentamer is dictated by a combination of hydrophobic, ionic and hydrogen-bonding interactions rendering it a stable and functional biological unit.

Active site

The active site of NdCld is located in the cavity of the C-terminal domain. This cavity, formed by the helices α2', α3', α4' and the β-sheet of the C-terminal domain (Fig. 2A) has a volume of ~1000 Å^3, and was found to accommodate one heme *b* moiety per subunit. Heme *b* is embedded in each cavity within a defined hydrophobic environment, where it is surrounded by Phe114, Lys141, Met157 Ala164, Leu201, Leu205, and Glu210 at the proximal site (Fig. 3A) and residues Leu122, Ile137, Ile139, Val171, Arg173, Phe186, Tyr188, and Phe190 at the distal site (Fig. 3B). The heme iron is coordinated by His160 from the α3' helix at the proximal side at distances ranging from 2.11 – 2.27 Å (Fig. 3A). Histidine 160 is hydrogen bonded to Glu210, the latter being conserved in Clds with known chlorite dismutation activity. This conserved H-bond increases the imidazolate character of proximal histidine thereby shifting the reduction potential of the heme iron to more negative values (similar to heme peroxidases that have a conserved proximal His-Asp pair). This might be important in stabilization of higher heme oxidation state(s) involved in chlorite dismutation.

Two heme *b* propionate groups form hydrogen bonds to the loop between β4 and α1', to the α2' helix, and to the β1' strand. In detail, one carboxylate group is within hydrogen bonding distances to the main chain amide nitrogen atoms of Thr109, Tyr110 and Val111, while the other carboxylate group interacts with the NE1 atom of Trp145 (Fig. 3C), suggested to act as the electron donor for the reduction of compound I to compound II (Lee et al., 2008). In the two wild-type NdCld structures reported here, cyanide or imidazole were found to coordinate the iron at the distal heme position (Fig. 4A and B). In each subunit cyanide anion is bound to the heme iron at an average distance of 1.91 ± 0.01 Å and is also within hydrogen bonding distance to a water molecule and a molecule of glycerol used for cryo protection. Imidazole coordination bond length is 2.18 ± 0.06 Å. In addition, one of the two ethylene glycol molecules was found within hydrogen bonding distance to the imidazole.

Imidazole was used to elute NdCld from the HisTrap columns and co-purified with the recombinant protein due to high affinity of imidazole to ferric heme. When imidazole-containing crystals were soaked in cyanide, the imidazole was displaced by cyanide at a concentration as low as 5 mM. Both compounds are known to serve as inhibiting ligands of Cld (this study; Hagedoorn et al., 2002; van Ginkel et al., 1996). Hence, the obtained crystal structures represent an inhibited state of Cld at the atomic level.

A close inspection of the distal heme site (Fig. 3B) showed that the only residue able to provide a positive charge, or to shift towards the active site in the presence of an anionic ligand,

8

is Arg173, suggesting that this residue plays an important role in the catalytic reaction of chlorite degradation. In the two NdCld wild-type structures, Arg173 is oriented away from the heme *b* iron and points towards the putative substrate entrance, indicating a possible role of this residue in gating the entrance to the active site. The side chain of Arg173 is stabilized in this position by hydrogen bonds between the guanidinium and hydroxyl groups of glycerol or ethylene glycol, coming from the cryo-solution, and makes no direct interactions to cyanide or imidazole (Fig. 4A and B). In thiocyanate-bound AoCld no cryoprotectant was found at the corresponding position and the arginine side chain points towards the heme iron (Supplemental Fig. 1), forming a hydrogen bond to thiocyanate *via* its guanidinium group.

NdCld mutants

A possible role of Arg (corresponding to position of Arg173 in NdCld) in the ClO_2^- degradation activity of Cld was suggested previously (de Geus et al., 2009; Ebihara et al., 2005), but until now no biochemical or structural evidence has been generated to support this notion. We hence prepared two mutants of NdCld (R173A and R173K) and firstly characterized them at the structural level. As expected, the overall structure of the NdCld mutants remained unchanged compared to wild-type protein (r.m.s.d. of ~ 0.6 Å over all Cα atoms; all subunits superimposed together). Furthermore, no significant changes in the conformation of the side-chains in the active site of NdCld were observed upon the arginine-to-alanine or arginine-to-lysine substitutions. In the R173K mutant, the Lys173 side-chain adopts a conformation similar to that of Arg183 of AoCld (Supplemental Fig. 1) and forms a hydrogen bond with the sulphate anion found in the active site cavity (Fig. 4D). Additionally, in both mutants a water molecule coordinates the iron at the distal part of heme *b* (at distances of 3.0 Å and 3.2 Å) (Fig. 4C and D), whereas none of these crystals contained a cryoprotectant molecule close to the active site.

Finally, the most pronounced structural differences between active sites of wild-type and R173K mutant reside in side-chain conformations of Arg173 and Lys173 (Supplemental Fig. 1), suggesting an embedded conformational flexibility for Arg173. This allows it to adopt a conformation similar to Lys173 mutant and to AoCld Arg183 in the active enzyme.

Steady-state activity

Table 2 summarizes the overall chlorite dismutase activity and apparent bimolecular rate constants of cyanide binding for wild-type NdCld and for the two variants R173A and R173K. Fig. 5A depicts the initial rate of O_2 release followed polarographically as a function of chlorite concentration. An apparent saturation in the initial rate with increasing chlorite concentration is obvious. Data could be best fitted only by a double rectangular hyperbola equation (for better clarity see semilogarithmic plot in inset to Fig. 5A). With increasing chlorite concentrations irreversible inactivation of the enzyme occurred as is evident also by inspection of individual time traces depicted in Fig. 5B. Thus, it was important to (i) use only the initial linear phase for rate calculation and to (ii) deduce Michaelis-Menten parameters from a set of chlorite concentrations with $[ClO_2^-] < 1$ mM, which reflect physiological conditions. This also explains why the Michaelis-Menten parameters of NdCld determined here differ from previously reported values (Maixner et al., 2008). In the previous study, the assays were conducted using much higher chlorite concentrations (up to 65 mM), which are unlikely to occur *in situ*. Thus, we regard the values measured here to be more accurate in terms of enzyme activity under natural conditions. For wild-type Cld K_M and k_{cat} were determined to be 58 µM and 35 s^{-1}, resulting in a

9

catalytic efficiency (k_{cat}/K_M) of 6.0×10^5 M^{-1} s^{-1}. For comparison, the catalytic efficiency and k_{cat} of Cld from *Azospira oryzae* GR-1 (with the Arg pointing to the heme in the crystal structure) were 11 times and 34 times higher compared to the kinetic parameters of NdCld (van Ginkel et al., 1996).

In principle, both Arg mutants showed the same behaviour as wild-type NdCld, exhibiting kinetics of dioxygen production that could also be fitted best with a double rectangular hyperbola. Compared to the wild-type protein, in both variants the calculated K_M values increased, whereas the k_{cat} values decreased. As a consequence in both mutated enzymes the catalytic efficiency of chlorite dismutation was diminished (Table 2).

Kinetics of cyanide binding and dissociation

Figures 5C-E show the spectral changes of ferric NdCld upon mixing with cyanide. This ligand converts the high-spin ($S = 5/2$) iron state to the low-spin ($S = 1/2$) state, thereby shifting the Soret peak from 410 to 422 nm with an intermediate at 418 nm. The first spectral transition from 410 nm to 418 nm showed a clear isosbestic point at 415 nm, whereas the second spectral transition from 418 to 422 nm did not show any isosbestic point (Fig. 5C). For both mutants cyanide binding was monophasic with a shift of the Soret maximum from 410 nm to 422 nm and a clear isosbestic point at 418 nm (not shown). By using the stopped-flow apparatus, cyanide binding was followed at 408 nm for all three recombinant proteins. In presence of excess cyanide, pseudo-first-order rate constants, $k_{obs(1)}$ and $k_{obs(2)}$, could be obtained from double-exponential fits for wild-type NdCld (Fig. 5C). In contrast, in both mutants the reactions were monophasic and k_{obs} values could be obtained from single-exponential fits. The apparent second-order rate constant for cyanide binding (k_{on}) was calculated from the slope of the linear plot of $k_{obs(1)}$, or k_{obs} *versus* the cyanide concentration [$k_{obs} = k_{on}$[HCN] + k_{off}] (Fig. 5D). Cyanide binding to wild-type NdCld [$(1.57 \pm 0.07) \times 10^6$ M^{-1} s^{-1} at 25 °C] was about 460 times faster compared to R173A [$(3.43 \pm 0.07) \times 10^3$ M^{-1} s^{-1}] and 970 times faster than R173K mutant [$(1.62 \pm 0.01) \times 10^3$ M^{-1} s^{-}]. From the intercept of the linear plots (Fig. 5D) the dissociation rate constants (k_{off}) were obtained, allowing the calculation of the dissociation constants (K_D) of the cyanide complexes from the k_{off}/k_{on} ratios. The K_D-value for wild-type NdCld is 3.6 µM and increased by factors of 40 to 50 times for the R173A (146 µM) and R173K (185 µM) mutant, respectively.

Comparative amino acid sequence analysis of Clds and Cld-like proteins

A multiple sequence alignment of 174 Clds and Cld-like proteins was created by using as template a structure-based sequence alignment of the five crystallized proteins from this superfamily. This alignment allowed us to investigate the conservation of the heme environment in these enzymes from different main microbial lineages of descent (phyla), by analysis of residues located within less than 4 Å distance from the cofactor. The high degree of conservation of heme-binding residues (Supplemental Fig. 2) strongly suggests that all Clds and Cld-like enzymes contain non-covalently bound protoporphyrin IX or heme *b* as cofactor, whereas no analyzed sequence contained the characteristic binding motif (Bowman and Bren, 2008) of covalently bound heme *c* (CXXCH). The proximal heme *b* ligand His160 is highly conserved in the Clds and Cld-like proteins, with the only exception of the respective proteins from *Leptospirillum*, a genus belonging to the *Nitrospirae* phylum. Similarly, functional importance of

10

the Trp145 residue as electron donor is reflected in its full conservation throughout the entire Cld superfamily. The catalytically important residue Arg173 of NdCld, and the homologous residues in other Clds and Cld-like proteins, deserve special attention. At this position, Arg is conserved in the biochemically characterized and highly active Clds of *Ca. N. defluvii* and *A. oryzae* (Supplemental Fig. 2) and in other PCRBs. In contrast, other residues are found in most but not all of the yet uncharacterized Cld-like proteins and in the enzyme from *T. thermophilus*, which has only a very weak Cld activity (Supplemental Fig. 2). This conservation pattern supports the proposed functional role of Arg at this position for efficient chlorite degradation, and is consistent with the low catalytic efficiency of the NdCld mutants R173A and R173K.

DISCUSSION

Functional NdCld is a pentamer

Cld-like enzymes from different sources have been described as tetramers in their native state (Coates et al., 1999; Mehboob et al., 2009; Streit and DuBois, 2008; van Ginkel et al., 1996). In many of these studies classical biochemical approaches such as SEC were used to determine the molecular weight of the protein. In contrast, in most crystal structures of Cld-like enzymes deposited in the PDB database, the proteins were found to exhibit a pentameric organization. The only exception is the structure of AoCld that forms a hexamer in the crystal, although it is pentameric in solution (de Geus et al., 2009). In agreement with other crystal structures we found NdCld to form a pentamer in the crystal, as well as in solution, implying that functional NdCld has a pentameric composition. Thus, since SEC might fail to estimate the real molecular weight, one could speculate that Cld-like enzymes identified as tetramers by SEC actually were pentamers. On the other hand, subunits of Cld-like enzymes may adopt different oligomerization states as was shown for AoCld. It remains to be shown in future experiments whether Cld-like enzymes adopt exclusively a pentameric arrangement in solution.

The active site pocket of NdCld is accessible from the outside of the pentameric ring through two channels one of which is shallower and was in all structures occupied by a sulfate or a phosphate ion. The second opening, observed also in the AoCld structure (de Geus et al., 2009), is deeper with a much higher positive surface potential, suggesting that it could serve both as substrate entry and product exit site (Fig. 6A). In case of R173K mutant the access to the active site through this channel is partially blocked compared to wild type (Fig. 6B and C), while the R173A mutant exhibits notably smaller reduced positive surface potential (Fig. 6D). The entrance to the active site from the central channel of the holoenzyme, which was reported for the hexameric AoCld (de Geus et al., 2009), was not observed for the pentameric NdCld.

Structural conservation of Cld fold

Comparisons of known subunit structures of Cld-like enzymes revealed a pronounced structural conservation of the Cld subunit fold (1.72 Å r.m.s.d. over 180 structurally equivalent Cα atoms of all subunits in all available structures), contrasting low average amino acid sequence identity among compared proteins (Supplemental Fig. 2 and Maixner et al., 2008). The main structural differences lay in conformation of the loop region between β4 and α1' which is involved in binding of propionate groups of heme *b*. Subunits of NdCld and AoCld which display the most similar conformation of this region accommodate heme *b* in their active sites,

11

while the prosthetic group is absent in the structures of other Cld-like enzymes (*T. acidophilum*, *G. stearothermophilus* and *T. thermophilus)*. Moreover, in the case of the Cld-like protein from *T. thermophilus* HB8, neither the addition of the heme precursor 5-aminolevulinic acid, nor the reconstitution of protein with iron protoporphyrin IX resulted in a protein stoichiometrically loaded with heme (Ebihara et al., 2005). Further research should clarify whether the observed lack of heme in these proteins was an experimental artefact. The heme ligands, which are conserved in almost all of the analyzed sequences including those from *T. acidophilum*, *G. stearothermophilus* and *T. thermophilus* (Supplemental Fig. 2), indicate that the binding of heme cofactor is a common feature of the whole enzyme superfamily.

Role of Arg173 in chlorite degradation

A role of the guanidinium group of the arginine at position 173 (NdCld numbering) in substrate positioning and activation during chlorite degradation by Cld has been suggested earlier (de Geus et al., 2009). Here, this hypothesis was investigated for the first time in a comprehensive structural, biochemical and bioinformatics approach. Exchange of Arg173 by either Ala or Lys significantly reduced the association rate of cyanide. This low-spin ligand is a useful probe to test the accessibility of the heme cavity and the role of distal residues in its positioning and stabilization of the resulting complex. Similar to hydrogen peroxide in catalases and peroxidases or to chlorite in Cld, cyanide binds in its anionic form and the cyanide complex mimics to some extent higher oxidation states like oxoiron(IV) species (i.e. compound I-& II). The calculated apparent bimolecular k_{on} rates of R173A and R173K were three orders of magnitude lower than that of the wild-type protein. In addition, the dissociation rates were diminished. These data clearly suggest a prominent role of Arg173 in NdCld in low-spin ligand binding to the heme iron.

Both mutants R173A and R173K still exhibited chlorite-degrading activity. Exchange of Arg173 lowered the affinity towards ClO_2^- (Table 2) and the turnover rates of R173A and R173K were diminished. These data suggest that Arg173 is important but not absolutely essential for catalysis. Comparison of wild-type NdCld with the mutants R173K and R173A revealed a remarkable conformational conservation between these structures, clearly demonstrating that Arg173 does not play a key role in the stabilization of the protein structure. Namely, our structural data suggest that Arg173 is flexible and might be involved in the regulation of substrate or ligand uptake and the stabilization of heme complexes or transient reaction intermediates. Such a primary role of Arg173 is in agreement with its properties (charge and shape) and its position at the distal heme side, not far away from the reaction centre and still close to the entrance of the substrate channel. Upon its exchange both the cyanide association rate is reduced and the K_M for chlorite is increased compared to the wild-type protein. Interestingly, the K_M value of R173K was found to be even higher than that of R173A. This suggests a restricted substrate access to the active site in presence of lysine, which is indicated also by the position of lysine *versus* arginine as found in the crystal structures of R173K and wild-type NdCld, respectively (Fig. 4B and D, Supplemental Fig. 1). However, while a significant decrease in k_{cat} was observed for R173A, the k_{cat} of R173K resembled the value of wild-type NdCld. Thus, it seems that in presence of sufficient amount of substrate, lysine may efficiently substitute for arginine, suggesting that a positive charge at this site is important for effective substrate processing. In addition, a positively charged residue will promote the entry of chlorite to the heme cavity through electrostatic attraction. Accordingly, considerable differences

12

in the surface potential at the substrate entrance of the wild-type NdCld and mutant R173A were found (Fig. 6).

Based on the mechanism of the Cld reaction proposed by Lee et al. (2008), compound I (an oxoiron(IV) porphyrin π-cation radical) and ClO⁻ are formed during the reaction of chlorite with the ferric enzyme. Anionic hypochlorite must remain close to the active site to perform its nucleophilic attack on oxoiron(IV). Since removing the positive charge at position 173 causes a reduction of the overall dismutation rate, it is reasonable to assume that Arg173 helps in addition to stabilize the transient compound I – hypochlorite complex. Bleaching of the heme and simultaneous loss of activity upon exposure of Clds to chlorite was observed in several studies and was attributed to the oxidation of the heme (Stenklo et al., 2001; van Ginkel et al., 1996). Consistent with a stabilizing role of the positively charged Arg173 one might also speculate that released hypochlorite contributes to heme bleaching. This fits with the observation that with increasing amounts of chlorite, bleaching of the NdCld heme was faster in R173A compared to both R173K and the wild-type protein (data not shown).

A second anion binding site that could accommodate ClO⁻ for further transformation has been identified near the active site of AoCld. At this anion binding site, hydrogen carbonate was found to coordinate the strictly conserved residue Lys114 (de Geus et al., 2009). However, as the distance between carbonate and the heme iron in AoCld is 10-12 Å, it appears unlikely that intermediates can be stabilized at this position during the fast chlorite degradation reaction. Thus, although NdCld has a similar anion binding site, we propose that the guanidinium group of Arg173 mainly contributes to keeping ClO⁻ in place for formation of the O-O bond in the second step of chlorite dismutation.

Active site conservation and evolution of Cld

The results of our combined structural and bioinformatics analyses indicate that the second ferredoxin-like domain of Cld subunits has evolved in a protoporphyrin IX or heme *b* binding moiety in all Clds and Cld-like enzymes. The overall high degree of structural conservation allowed us to carry out a phylogenetic analysis, using a structure-based refined alignment, on a set of selected Cld and Cld-like protein sequences. This analysis revealed a similar phylogeny of this enzyme superfamily as discussed in a previous study (Maixner et al., 2008). The reconstructed phylogenetic tree (Fig. 7) shows the wide distribution of Cld-like proteins across the *Bacteria* and *Archaea* with a pronounced grouping of proteins of organisms from the same major phylogenetic lineage (phylum). However, the topology of the protein tree only partly reflects 16S rRNA-based phylogeny. For example, NdCld is closely related to the Clds of proteobacterial PCRB although *Proteobacteria* and *Nitrospirae* are related only at the lowest interphylum level. This observation suggests a functional clustering of Clds and Cld-like proteins in the tree, indicating that closely related enzymes play similar functional roles in the respective organisms. This hypothesis is supported by mapping the homologous amino acid residues at position 173 (NdCld numbering) on the phylogenetic lineages. All validated Clds from known PCRBs and from *Nitrospira*, and yet uncharacterized Cld-like proteins from two bacterial phyla (*Proteobacteria* and *Cyanobacteria*), contain Arg at this position and form a major branch of the enzyme superfamily (Fig. 7). Our experimental results with wild-type NdCld and the mutants R173A and R173K show that a high catalytic efficiency of Cld depends on this Arg residue, which thus may serve as a signature of efficient Clds. On this basis, it is tempting to speculate

13

that the yet uncharacterized Cld-like proteins containing the critical Arg might represent previously overlooked efficient Clds from organisms that were not known to possess this activity. This hypothesis should be verified in future studies, as it could have strong implications for dissecting the composition and structure-function relationships of (per)chlorate-removing microbial consortia. In contrast, in the Cld-like enzyme of *T. thermophilus* the Arg is replaced by Gln, which is also present in a large number of related proteins from other bacterial and archaeal phyla (Fig. 7). As the Cld activity of the protein from *Thermus* is very low (Ebihara et al., 2005), it appears that the Cld-like proteins bearing the Gln residue may share a biological function, which is not chlorite degradation *in vivo*. Likewise, we assume that the Cld-like proteins of most *Actinobacteria*, which host Ala residues, the archaeal enzymes containing Ser, and the forms bearing other residues at position 173 (NdCld numbering) (Fig. 7) represent different functional groups of enzymes. Even though these proteins might still have a detectable but low Cld activity, resembling our NdCld mutants, a primary role of such enzymes in chlorite detoxification appears unlikely. Chlorite is highly reactive and has toxic effects on bacteria already at low concentrations of 10-20 µM (van Wijk et al., 1998). Hence, the high substrate affinity and turnover rates of *bona fide* Clds (containing the Arg residue) should be essential for PCRBs or other organisms, which are exposed to chlorite and are protected by this enzyme.

The presence of both a phylogenetic signal and functional clustering in the protein tree implies that starting from a common ancient motif, structural and functional variants of Cld-like proteins evolved in the different phyla. The presence of similar enzymes in phylogenetically distant organisms can be explained by horizontal gene transfer or, alternatively, by convergent or parallel evolution driven by similar selective pressures. Interestingly, no larger group of organisms (except the PCRB) containing similar Cld-like proteins has obvious ecophysiological traits in common, which could point at possible functions of the respective enzymes. Hence, the biological roles of most proteins in the Cld superfamily, including those which were found in important pathogenic bacteria, remain mysterious, and their experimental clarification will pose an interesting challenge for future research.

Conclusions

In this study we provide new data that underline unique structural and functional features of distal Arg173 in Cld. This residue, which is conserved in all sequenced Cld-like enzymes with a confirmed high chlorite dismutation activity, controls and supports ligand and substrate access and binding and keeps transiently formed hypochlorite in place for nucleophilic attack of ferryl oxygen in compound I. These findings are in agreement with proposed catalytic mechanism(s) (Lee et al., 2008; Streit and DuBois, 2008). The presence of this Arg residue in yet uncharacterized Cld-like proteins indicates a broader distribution of Clds, and the ability to detoxify chlorite, among microbes than was previously anticipated. Furthermore, based on our structural and bioinformatics analysis we postulate that all Clds and Cld-like enzymes do noncovalently bind protoporphyrin IX or heme *b*.

14

REFERENCES

Abrahams, J.P., S.K. Buchanan, M.J. Van Raaij, I.M. Fearnley, A.G. Leslie, and J.E. Walker, 1996. The structure of bovine F1-ATPase complexed with the peptide antibiotic efrapeptin. Proc Natl Acad Sci U S A 93: 9420-4.

Berry, E.A., and B.L. Trumpower, 1987. Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. Anal Biochem 161: 1-15.

Bowman, S.E., and K.L. Bren, 2008. The chemistry and biochemistry of heme c: functional bases for covalent attachment. Nat Prod Rep 25: 1118-30.

CCP4, 1994. The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst D50: 760-763.

Coates, J.D., and L.A. Achenbach, 2004. Microbial perchlorate reduction: rocket-fueled metabolism. Nat Rev Microbiol 2: 569-80.

Coates, J.D., U. Michaelidou, R.A. Bruce, S.M. O'Connor, J.N. Crespi, and L.A. Achenbach, 1999. Ubiquity and diversity of dissimilatory (per)chlorate-reducing bacteria. Appl Environ Microbiol 65: 5234-41.

Davis, I.W., L.W. Murray, J.S. Richardson, and D.C. Richardson, 2004. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic Acids Res 32: W615-9.

de Geus, D.C., E.A. Thomassen, C.L. van der Feltz, and J.P. Abrahams, 2008. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of chlorite dismutase: a detoxifying enzyme producing molecular oxygen. Acta Crystallogr Sect F Struct Biol Cryst Commun 64: 730-2.

de Geus, D.C., E.A. Thomassen, P.L. Hagedoorn, N.S. Pannu, E. van Duijn, and J.P. Abrahams, 2009. Crystal structure of chlorite dismutase, a detoxifying enzyme producing molecular oxygen. J Mol Biol 387: 192-206.

DeLano, W.L., 2000. The PyMOL Molecular Graphics System on World Wide Web http://www.pymol.org.

Ebihara, A., A. Okamoto, Y. Kousumi, H. Yamamoto, R. Masui, N. Ueyama, S. Yokoyama, and S. Kuramitsu, 2005. Structure-based functional identification of a novel heme-binding protein from Thermus thermophilus HB8. J Struct Funct Genomics 6: 21-32.

Emsley, P., and K. Cowtan, 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126-32.

Goblirsch, B.R., B.R. Streit, J.L. Dubois, and C.M. Wilmot, 2010. Structural features promoting dioxygen production by Dechloromonas aromatica chlorite dismutase. J Biol Inorg Chem.

Hagedoorn, P.L., D.C. De Geus, and W.R. Hagen, 2002. Spectroscopic characterization and ligand-binding properties of chlorite dismutase from the chlorate respiring bacterial strain GR-1. Eur J Biochem 269: 4905-11.

Kabsch, W., 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Cryst 26: 795-800.

Krissinel, E., and K. Henrick, 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr 60: 2256-68.

Lee, A.Q., B.R. Streit, M.J. Zdilla, M.M. Abu-Omar, and J.L. DuBois, 2008. Mechanism of and exquisite selectivity for O-O bond formation by the heme-dependent chlorite dismutase. Proc Natl Acad Sci U S A 105: 15654-9.

15

Letunic, I., and P. Bork, 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23: 127-8.

Maixner, F., M. Wagner, S. Lucker, E. Pelletier, S. Schmitz-Esser, K. Hace, E. Spieck, R. Konrat, D. Le Paslier, and H. Daims, 2008. Environmental genomics reveals a functional chlorite dismutase in the nitrite-oxidizing bacterium 'Candidatus Nitrospira defluvii'. Environ Microbiol 10: 3043-56.

Mehboob, F., A.F. Wolterink, A.J. Vermeulen, B. Jiang, P.L. Hagedoorn, A.J. Stams, and S.W. Kengen, 2009. Purification and characterization of a chlorite dismutase from Pseudomonas chloritidismutans. FEMS Microbiol Lett 293: 115-21.

Morris, R.J., A. Perrakis, and V.S. Lamzin, 2003. ARP/wARP and automatic interpretation of protein electron density maps. Methods Enzymol 374: 229-44.

Murshudov, G.N., A.A. Vagin, and E.J. Dodson, 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240-55.

Spieck, E., C. Hartwig, I. McCormack, F. Maixner, M. Wagner, A. Lipski, and H. Daims, 2006. Selective enrichment and molecular characterization of a previously uncultured Nitrospira-like bacterium from activated sludge. Environ Microbiol 8: 405-15.

Stenklo, K., H.D. Thorell, H. Bergius, R. Aasa, and T. Nilsson, 2001. Chlorite dismutase from Ideonella dechloratans. J Biol Inorg Chem 6: 601-7.

Streit, B.R., and J.L. DuBois, 2008. Chemical and steady-state kinetic analyses of a heterologously expressed heme dependent chlorite dismutase. Biochemistry 47: 5271-80.

Ueno, H., K. Oishi, Y. Sayato, and K. Nakamuro, 2000. Oxidative cell damage in Kat-sod assay of oxyhalides as inorganic disinfection by-products and their occurrence by ozonation. Arch Environ Contam Toxicol 38: 1-6.

Uson, I., and G.M. Sheldrick, 1999. Advances in direct methods for protein crystallography. Curr Opin Struct Biol 9: 643-8.

Vagin, A., and A. Teplyakov, 1997. MOLREP: an Automated Program for Molecular Replacement. J Appl Cryst 30: 1022-1025.

van Ginkel, C.G., G.B. Rikken, A.G. Kroon, and S.W. Kengen, 1996. Purification and characterization of chlorite dismutase: a novel oxygen-generating enzyme. Arch Microbiol 166: 321-6.

van Wijk, D.J., S.G. Kroon, and I.C. Garttener-Arends, 1998. Toxicity of chlorate and chlorite to selected species of algae, bacteria, and fungi. Ecotoxicol Environ Saf 40: 206-11.

Vonrhein, C., E. Blanc, P. Roversi, and G. Bricogne, 2007. Automated structure solution with autoSHARP. Methods Mol Biol 364: 215-30.

FOOTNOTES

[$]The abbreviations used are: Cld, chlorite dismutase; PCRB, perchlorate-reducing bacteria; NOB, nitrite-oxidizing bacteria; NdCld, chlorite dismutase of "*Candidatus* Nitrospira defluvii"; SEC, size exclusion chromatography; AoCld, chlorite dismutase of *Azospira oryzae* strain GR-1; DaCld, chlorite dismutase of *Dechloromonas aromatica*; R173K, mutant of NdCld where Arg at position 173 was substituted by Lys; R173A, mutant of NdCld where Arg at position 173 was substituted by Ala; r.m.s.d., root mean square deviation.

16

ACKNOWLEDGMENTS

17

FIGURE LEGENDS

FIGURE 1. **Crystal structure of Cld from *Ca*. Nitrospira defluvii**. Ribbon representation of the NdCld structure. Monomers are shown in a unique color. The heme group is shown as a blue sticks model in all monomers. The iron is displayed as a red sphere. View of the pentameric ring-like arrangement of NdCld subumits seen from the proximal site of the heme. All graphic representations were generated using PyMol (DeLano, 2000).

FIGURE 2. **Structure of a NdCld subunit**. (A) The α helices and β strands are labeled according to the ferredoxin fold. The NdCld monomer consists of two ferredoxin-like domains, therefore secondary structure elements in the C-terminal domain follow the labeling of the N-terminal domain, with prime mark used as distinguisher. The α helices and β strands are shown in red and yellow, respectively. The heme is shown in blue, with iron shown as a red sphere. (B) The N-terminal domain (green) superimposed with the C-terminal domain (red) of the NdCld subunit. Secondary structure elements of N-terminal domain are labeled.

FIGURE 3. **Heme interactions with NdCld.** The figure shows residues interacting with heme in imidazole bound NdCld on its proximal (A), and distal part (B), as well as the interactions between heme propionates and protein (C). The carbon, oxygen, and nitrogen atoms are shown in green, red, and blue, respectively. The sulfur atom of the sulfate anion (*SO4*) is shown in orange. The heme iron and water molecules are shown as red spheres. Hydrogen bonds are drawn as dashed lines. Interactions of heme iron with His160 (*H160*) and imidazole (*IMD*) shown in (A) and (B), respectively, are also indicated by dashed lines.

FIGURE 4. **Active site in wild-type and mutant NdClds.** (A) Imidazole (*IMD*) was found to interact with the heme iron at the distal site in the wild-type NdCld. (B) After soaking of wild-type NdCld crystals with KCN, imidazole was effectively replaced by cyanide (*CN*). Water molecules (*Wat*) interact with the heme iron at the distal position in the NdCld R173A (C) and in the NdCld R173K structure (D). Molecules of ethylene glycol (*EDO*), glycerol (*GOL*), or sulfate anions (*SO4*) from cryo-solution (A and B) or reservoir (D) solutions were found in the active site. Atoms are colored as indicated in the legend of Fig. 3. The $|2F_o – F_c|$ electron density map was contoured at the 1.5 σ level (A). Note the different orientation of the distal Lys173 in the NdCld R173K mutant (D), compared to Arg173 of the NdCld wild-type shown in (A) and (B).

FIGURE 5. **Monitoring of dioxygen evolution and cyanide binding by wild-type NdCld.** (A) Plot of the initial rate (v_0) of molecular oxygen evolution as a function of chlorite concentration. Inset depicts the corresponding semi-logarithmic plot. Points represent averages of three measurements. In addition single- (grey line) and double- (black line) rectangular hyperbolar fits are shown. Conditions: 50 mM phosphate buffer pH 7.0, 20 nM NdCld, 30 °C. (B) Selected time traces with different chlorite concentrations. Conditions as in (A). (C) Spectral transition of the reaction of 2.5 µM Cld with 50 µM cyanide measured in the conventional stopped-flow mode. The first spectrum shows the ferric Cld in its high-spin state. The second spectrum was recorded 1.3 ms after mixing, with subsequent spectra at 3.8, 6.4, 11.5, 40, 109, and 9988 ms showing the formation of the low-spin cyanide complex. Arrows indicate changes of absorbance with time. Conditions: 50 mM phosphate buffer, pH 7, and 25 °C. The spectral transition for the reaction of

18

Cld with cyanide at 408 mn fits to a double-exponential equation. The first exponential term ($k_{\mathrm{obs}(1)}$) of the double-exponential equation was dependent on the cyanide concentration whereas the second exponential term ($k_{\mathrm{obs}(2)}$) was not (inset). (D) Dependence of $k_{\mathrm{obs}(1)}$ values from the cyanide concentration. The association rate constant was calculated from the slope and the dissociation rate constant from the intercept. Final enzyme concentration: 2 µM Cld in 50 mM phosphate buffer, pH 7. (E) Typical time trace at 408 nm with a double-exponential fit (2 µM Cld and 25 µM cyanide). Inset shows the plot of $k_{\mathrm{obs}(2)}$ against cyanide concentration.

FIGURE 6. **Molecular surface of the NdCld substrate entrance.** (A) The position and accessibility of the heme moiety in NdCld. The solvent accessible surface of one NdCld subunit is colored according to its electrostatic potential (blue for positive, red for negative). The semitransparent surface representations of other NdCld subunits forming the NdCld pentamer are shown in grey, with iron shown as a red sphere. Hemes are presented as green stick models. Detailed view into the active site chamber through the putative substrate entrance and product exit channel in the wild-type NdCld (B), NdCld mutant R173K (C), and NdCld mutant R173A (D). Note the partial blocking of the substrate entrance by the side chain of Lys173 (C), as well as differences in the entrance surface potential of both mutants (C and D) in comparison to the wild-type NdCld (B).

FIGURE 7. **Maximum likelihood tree based on amino acid sequences of Clds and Cld-like proteins.** Sequences marked with an asterix represent crystallized proteins. Colors depict the affiliations of the respective organisms to bacterial and archaeal phyla based on 16S rRNA phylogeny. Colored stripes define the residues at the position, which is homologous to Arg173 of NdCld, in the respective proteins. The circular tree was drawn by using the iTOL online tool (Letunic and Bork, 2007).

SUPPLEMENTAL FIGURE 1. **Structural comparison of Cld active sites.** The active sites of Cld from *Azospira oryzae* strain GR-1 (blue), NdCld (green), and of NdCld mutant R173K (red) were superimposed over equivalent Cα atom pairs. Differences in the active sites involve side-chain conformations of Arg173 and Lys173. Side chains of amino acid residues and the heme are shown as stick-models. The iron is displayed as a red sphere.

SUPPLEMENTAL FIGURE 2. **Alignment of selected chlorite dismutase (Cld) and Cld-like protein sequences.** Identifiers of proteins, whose structure has been experimentally determined, are printed bold. Conserved amino acid residues, which probably are involved in heme binding, are highlighted in orange. The asterisk indicates the alignment column of the characterized catalytic residue Arg173 in the Cld of *Ca*. N. defluvii. The homologous residues of the other Clds and Cld-like proteins at this position are highlighted according to the color scheme used in Fig. 7.

19

TABLES

TABLE 1.
Data collection and refinement statistics

	Data set 1	NdCld:IMD	NdCld:CN	NdCld R173A	NdCld R173K
Beamline	Home source	ID14-2 (ESRF)	ID14-2 (ESRF)	ID14-2 (ESRF)	ID14-1 (ESRF)
Wavelength (Å)	1.54	0.933	0.933	0.933	0.933
Resolution (Å)	40.7 - 2.40 (2.45 - 2.40)[a]	126.0-1.85 (1.9 - 1.85)	125.0-1.94 (1.99 - 1.94)	105.4 - 2.6 (2.67 - 2.60)	106.0 - 2.7 (2.77 - 2.70)
Space group	P3$_2$21	P3$_2$21	P3$_2$21	C2	C2
Unit cell (Å, º)	a = b = 145.29, c = 135.83	a = b = 145.67, c = 136.44	a = b = 145.49, c = 136.01	a = 139.74, b = 113.62, c = 120.52, β = 118.936	a = 137.41, b = 113.88, c = 119.87, β = 118.020
Molecules / a.u.	5	5	5	5	5
Unique reflections	126515 (7108)	268021 (18555)	121048 (17672)	92246	80774 (5070)
Completeness (%)	95.5 (93.3)	97.0 (90.7)	99.2 (99.8)	92.1 (79.9)	91.5 (78.2)
R$_{meas}$[b]		0.041 (0.351)	0.121 (0.638)	0.171 (0.459)	0.119 (0.474)
R$_{pim}$[c]	0.0211 (0.2663)				
Multipicity	24.1 (4.36)	5.09 (5.08)	6.4 (6.0)	5.8 (5.9)	4.76 (5.02)
I/sig(I)	14.5 (2.2)	24.4 (5.13)	12.3 (2.1)	8.46 (3.08)	8.67 (3.12)
No. of sites	29				
Phasing power[d]	0.474 (0.150)				
Figure of merit	0.232 (0.090)				
R$_{cryst}$[e]/ R$_{free}$[f]		0.177 / 0.211	0.206 / 0.254	0.210 / 0.249	0.206 / 0.244
R.m.s.d. bonds (Å)		0.013	0.015	0.011	0.011
R.m.s.d. angles (º)		1.411	1.455	1.199	1.256

[a]Values in parentheses are for the highest resolution shell.

[b]

$$R_{meas} = \frac{\sum_h \sqrt{\frac{n_h}{n_h-1}} \sum_i^{n_h} \left| \hat{I}_h - I_{h,i} \right|}{\sum_h \sum_i^{n_h} I_{h,i}} \text{ with } \hat{I}_h = \frac{1}{n_h} \sum_i^{n_h} I_{h,i}$$

[c]

$$R_{pim} = \sum_{hkl} \sqrt{\frac{1}{N-1}} \sum_i \left| I_i(hkl) - \overline{I(hkl)} \right| / \sum_{hkl} \sum_i I_i(hkl)$$

Where I (hkl) is the mean intensity of multiple I$_i$ (hkl) observations of the symmetry-related reflections, N is the redundancy, n$_h$ is the multiplicity, Î$_h$ is the average intensity and I$_{h,i}$ is the observed intensity.

[d]Anomalous phasing power: $(\Sigma |F_H(\text{imag})|^2 / \Sigma || \Delta F \pm_{PH}(\text{obs})| - |\Delta F \pm_{PH}(\text{calc})||^2)^{1/2}$ where $\Delta F \pm_{PH}$ is the structure factor difference between Bijvoet pairs and $F_H(\text{imag})$ is the imaginary component of the calculated structure factor contribution by the anomalously scattering atoms.

[e]$R_{cryst} = \Sigma |F_o - F_c| / \Sigma F_o$

[f]R$_{free}$ is the cross-validation R$_{factor}$ computed for the test set of reflections (5 %) which are omitted in the refinement process.

20

TABLE 2.

(A) Steady-state kinetic parameters for molecular oxygen evolution of wild-type NdCld and the variants R173A and R173K. (B) Pre-steady-state kinetic parameters for cyanide binding to wild-type NdCld and the variants R173A and R173K.

	wild-type	R173A	R173K
(A)			
K_M (µM)	58±37	107±54	818±250
k_{cat} (s^{-1})	35	7.1	21
k_{cat}/K_M (M^{-1} s^{-1})	6.0×10^5	6.6×10^4	2.5×10^4
(B)			
k_{on} (M^{-1} s^{-1})	2.57×10^6	3.43×10^3	1.62×10^3
k_{off} (s^{-1})	9.3	0.50	0.30
K_D (µM)	3.6	145.8	185.2

21

Fig. 1

A

B

Fig. 2

Fig. 3

Fig. 4

A

B

C

D

E
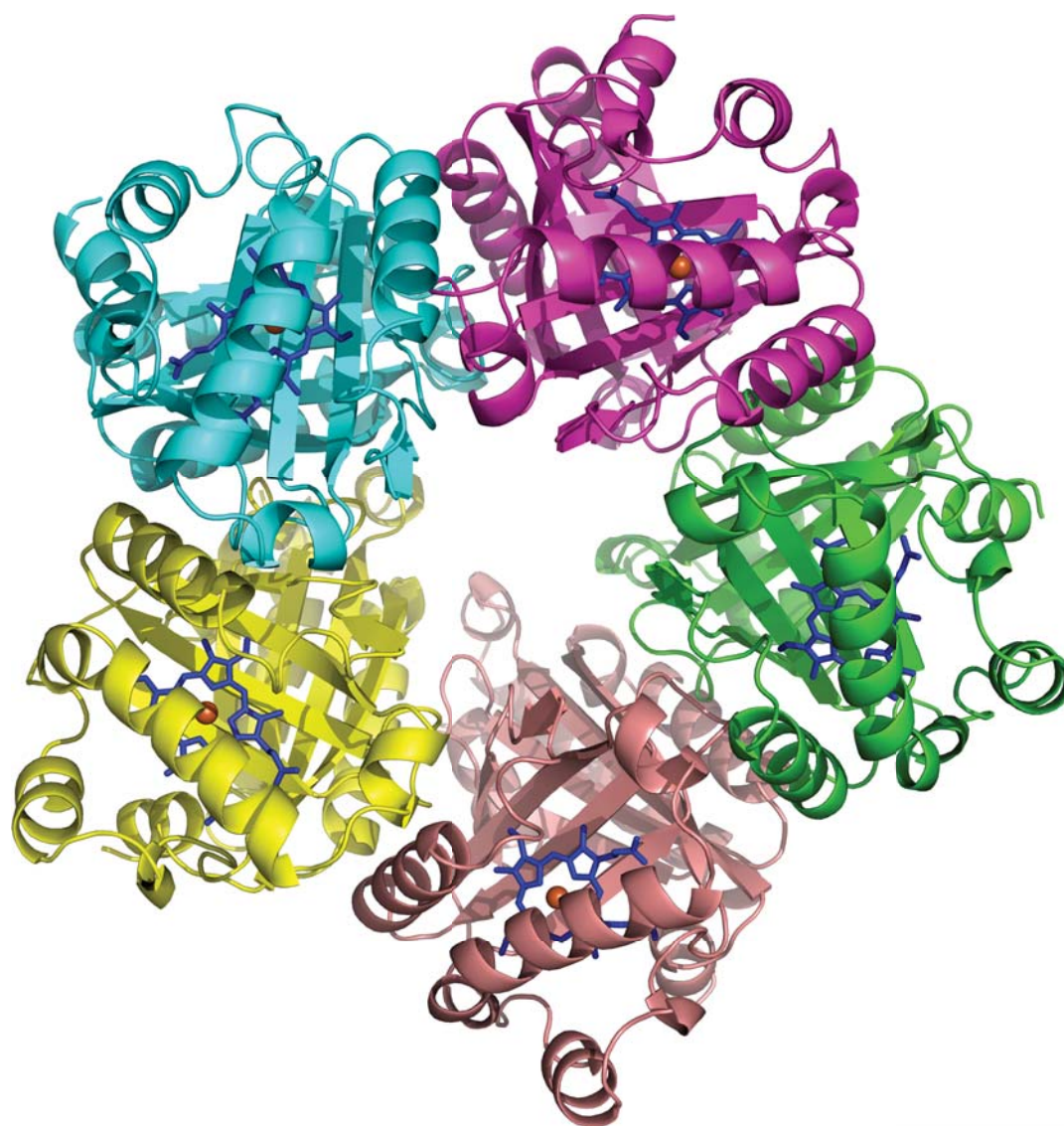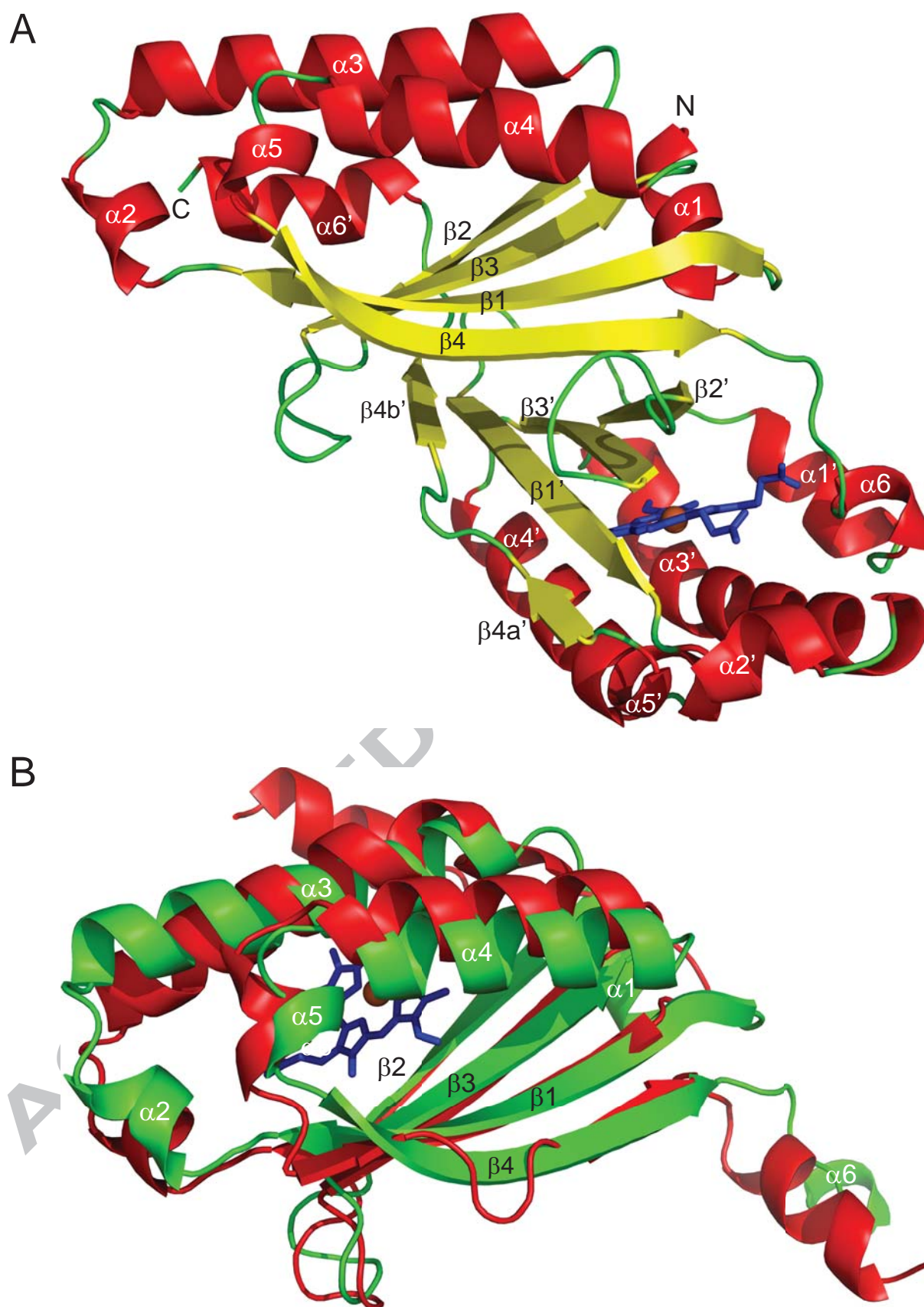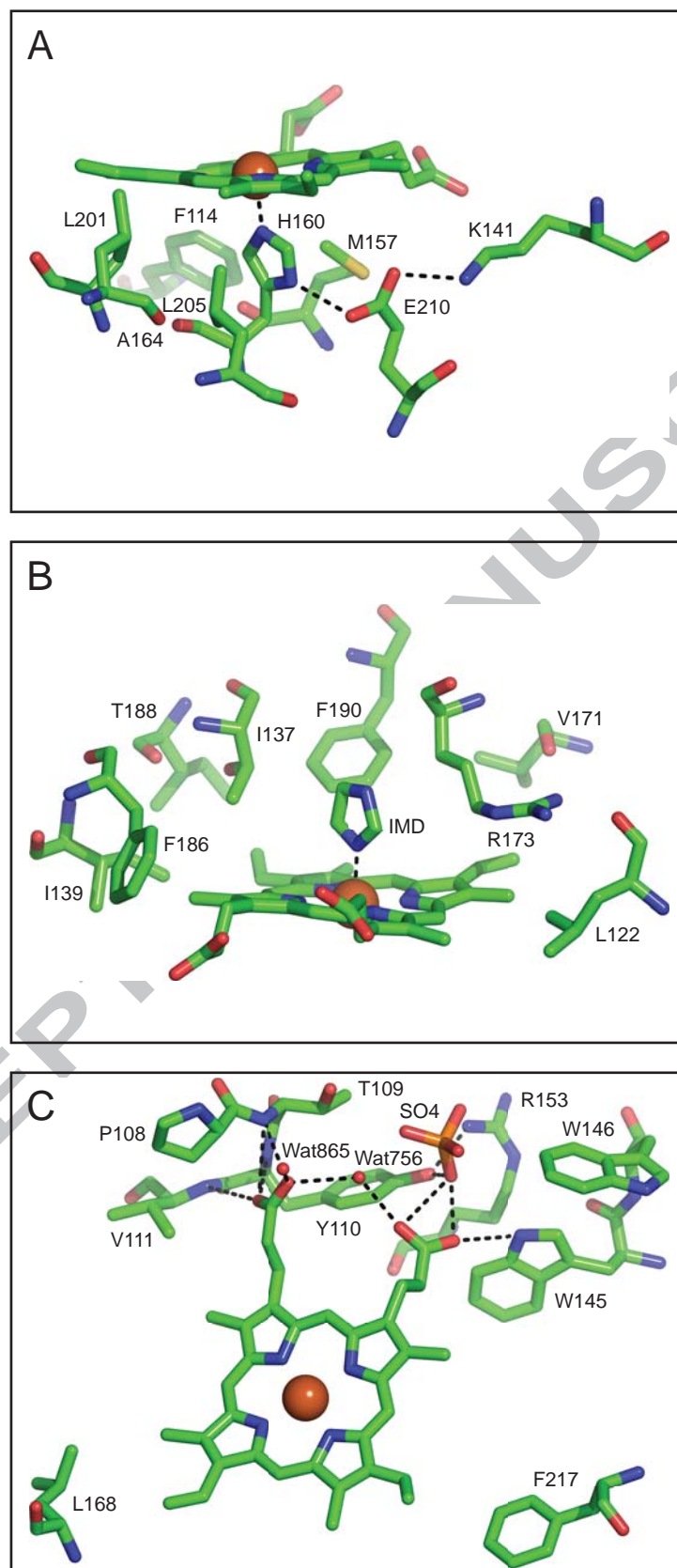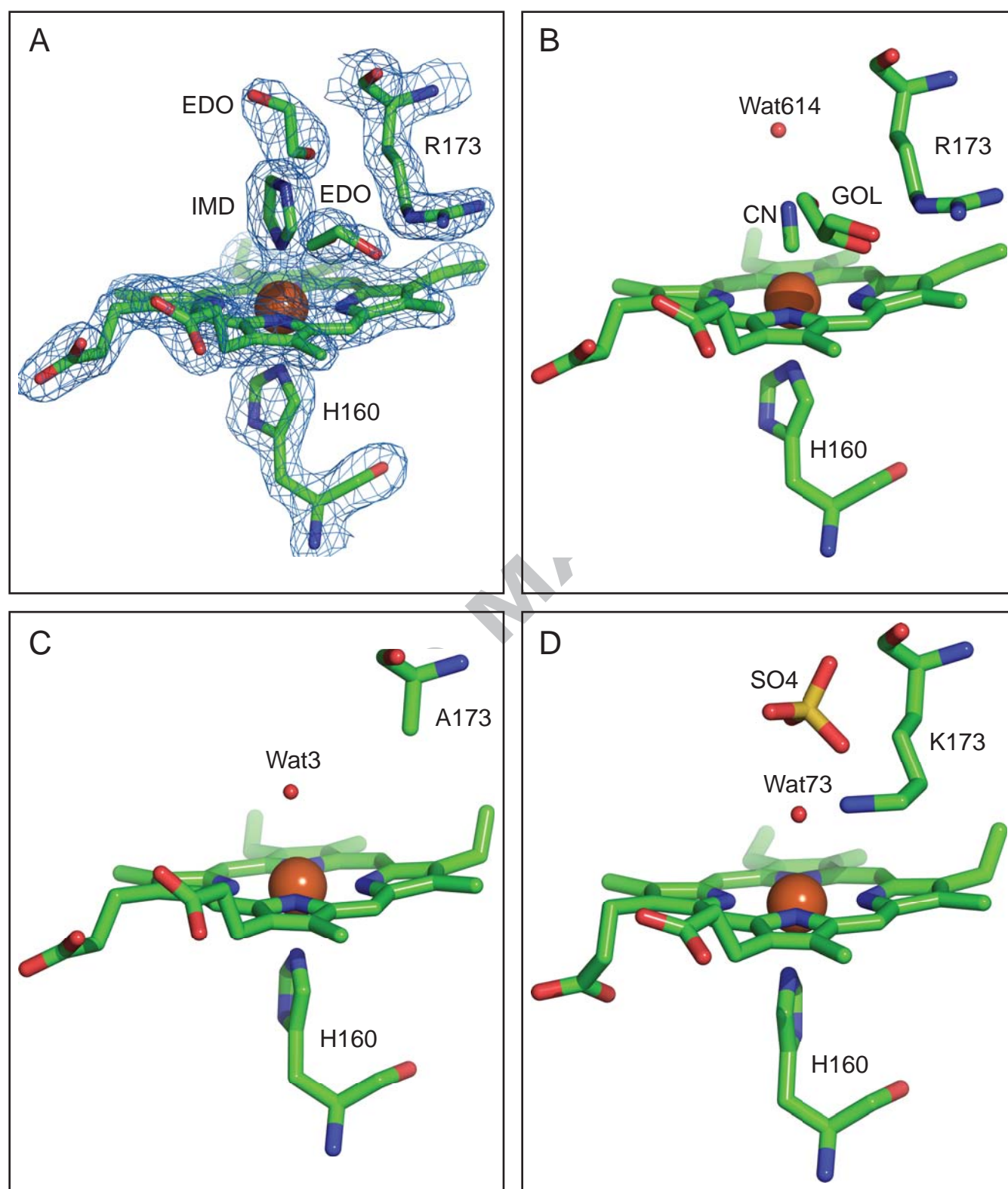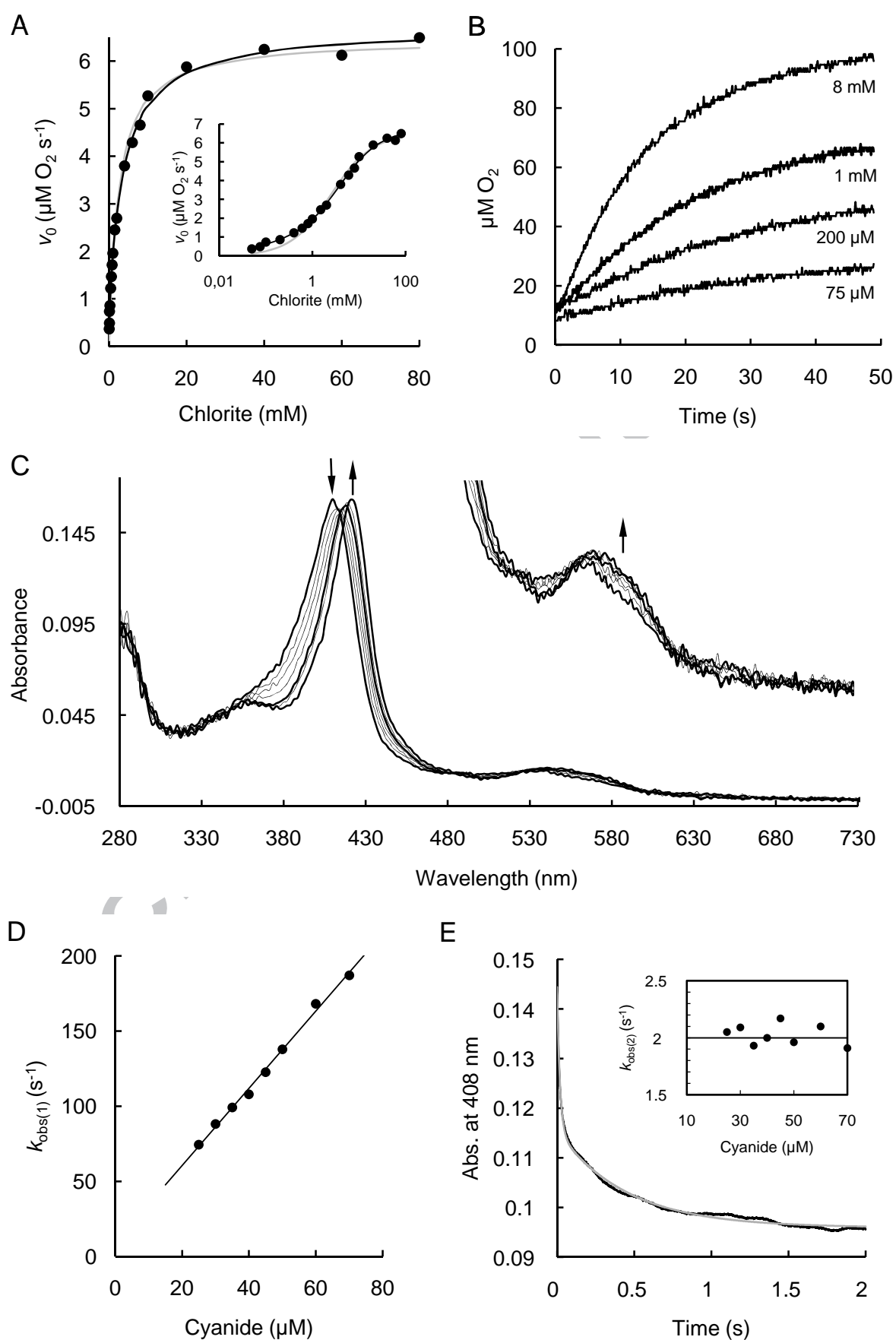
Fig. 5
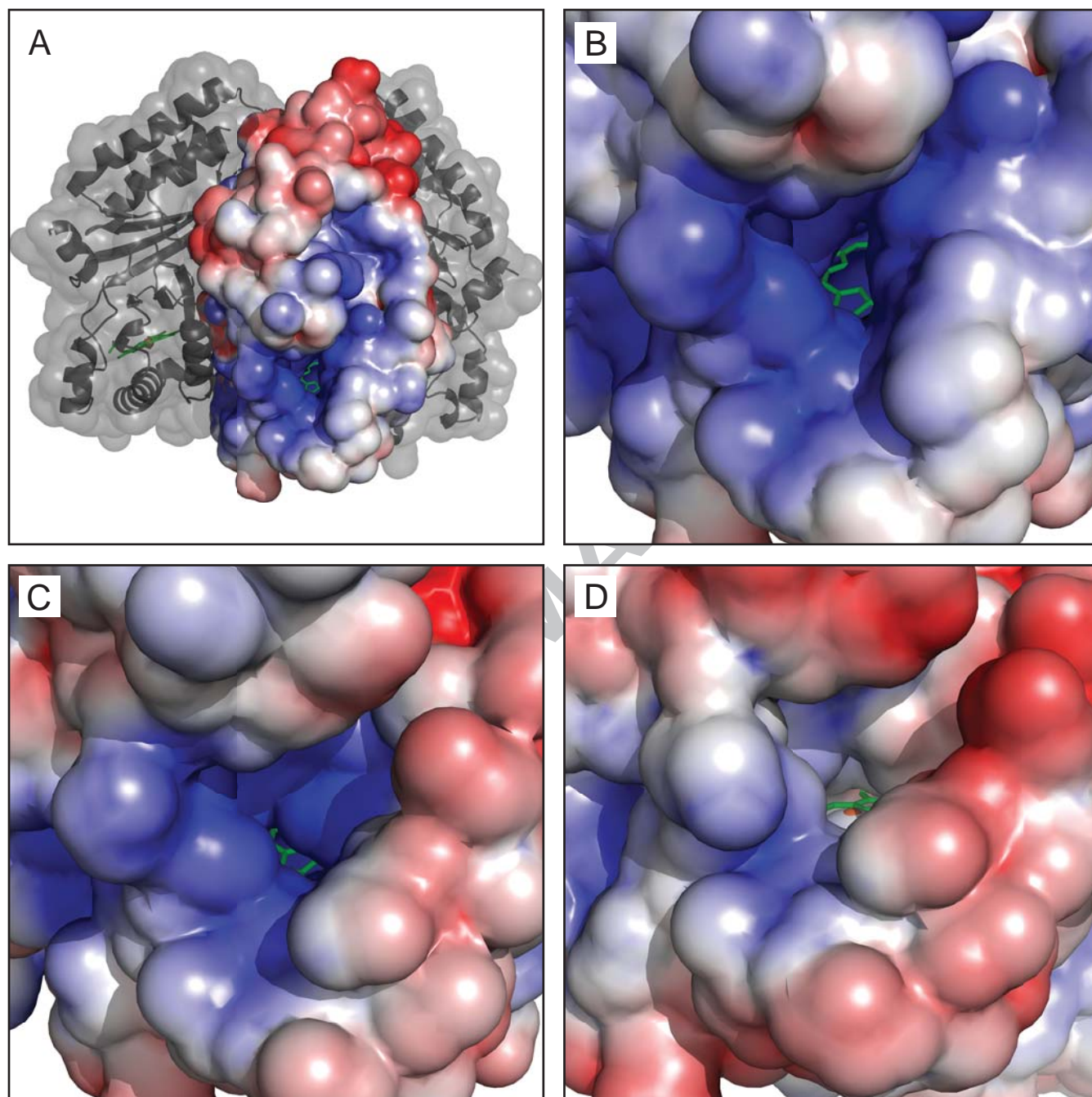
Fig. 6

Serine
Arginine
Leucine
Glutamine
Alanine
Others

Archaea
Nitrospirae
Proteobacteria
Cyanobacteria
Chloroflexi
Deinococcus-Thermus
Verrucomicrobia
Planctomycetes
Acidobacteria
Firmicutes
Actinobacteria

Fig. 7
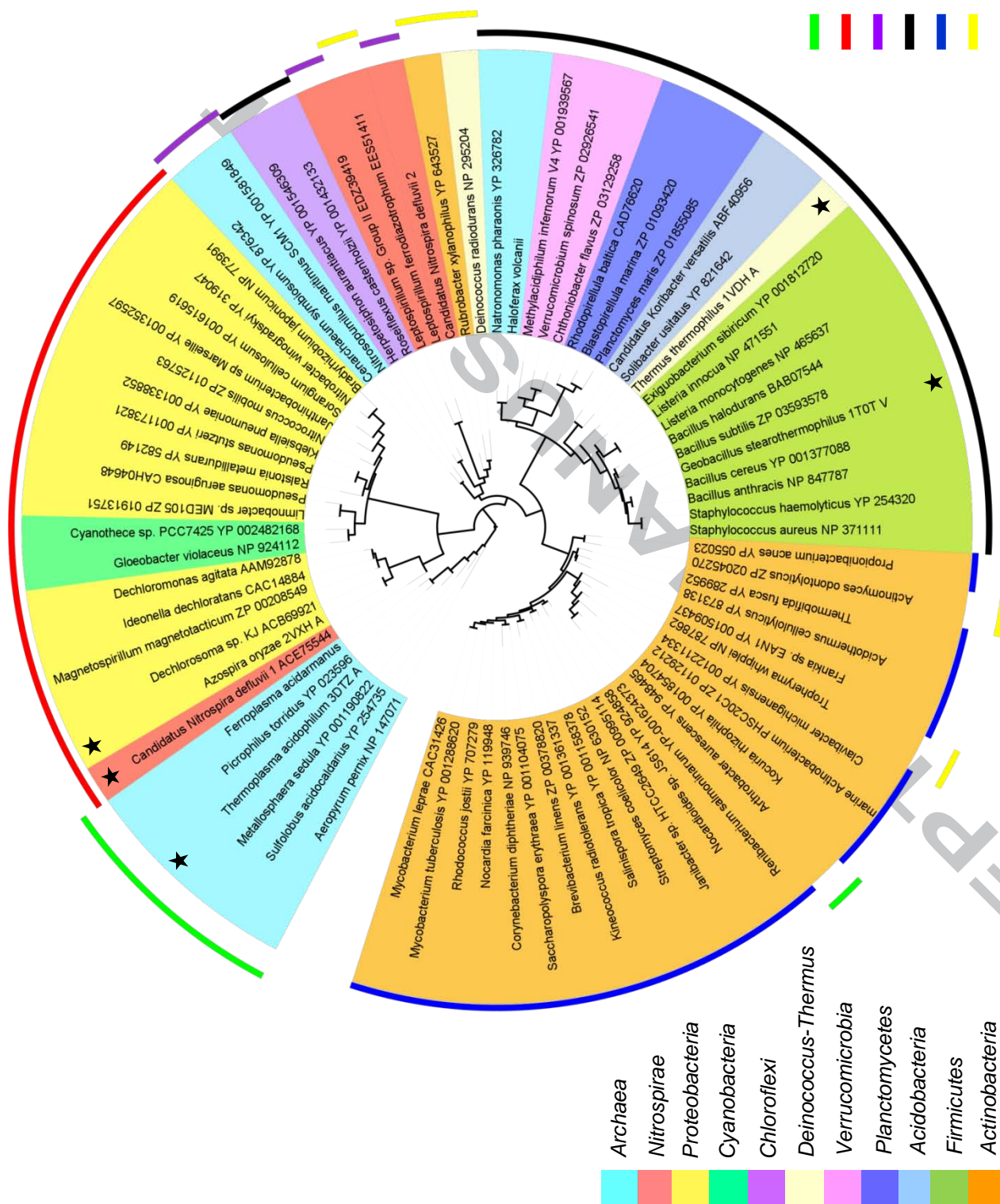
Chapter III

Unexpected Structural and Phylogenetic Diversity of Chlorite Dismutases: A Catalytically Efficient Dimeric Enzyme from *Nitrobacter winogradskyi*

Unexpected Structural and Phylogenetic Diversity of Chlorite Dismutases: A Catalytically Efficient Dimeric Enzyme from *Nitrobacter winogradskyi*

Georg Mlynek[1], Björn Sjöblom[1], Julius Kostan[1], Stephanie Füreder[2], Frank Maixner[2,5], Paul Georg Furtmüller[3], Christian Obinger[3], Michael Wagner[2], Holger Daims[2,*] and Kristina Djinović-Carugo[1,4,*]

[1]Department for Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, A-1030 Vienna, Austria; [2]Department of Microbial Ecology, Vienna Ecology Centre, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria; [3]Department of Chemistry, Division of Biochemistry, BOKU – University of Natural Resources and Applied Life Sciences, Muthgasse 18, Vienna, A-1190, Austria; [4]Department of Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, 1000 Ljubljana, Slovenia; [5]present address: Institute for Mummies and the Iceman, EURAC research, Viale Druso 1, 39100 Bolzano, Italy

Running title: A novel chlorite dismutase from *Nitrobacter winogradskyi*

Address correspondence to: Kristina Djinović-Carugo, Email: kristina.djinovic@univie.ac.at, Tel.: +43-1-4277-52203; Fax: +43-1-4277-9522; Holger Daims, Email: daims@microbial-ecology.net, Tel.: +43-1-4277-54392; Fax: +43-1-4277-54389

Subject Category: Microbial engineering

Abstract

Chlorite dismutase (Cld) is a unique heme enzyme catalyzing the conversion of ClO_2^- to Cl^- and O_2. Cld is usually found in (per)chlorate-reducing heterotrophic bacteria, but was recently identified also in a nitrite-oxidizing bacterium of the genus *Nitrospira*. Here we characterized a novel Cld-like protein from the chemolithoautotrophic nitrite oxidizer *Nitrobacter winogradsyki*, which is significantly smaller than all previously known chlorite dismutases. Its 3D crystal structure revealed a dimer of two identical subunits, which sharply contrasts the penta- or hexameric structures of other chlorite dismutases. Despite a truncated N-terminal domain in each monomer, this novel enzyme turned out to be a highly efficient chlorite dismutase, demonstrating a greater structural and phylogenetic diversity of these enzymes than previously known. Based on comparative analyses of Cld sequences and 3D structures, signature amino acid residues were identified that are useful to assess whether uncharacterized Cld-like proteins may have a high chlorite-dismutating activity. Interestingly, proteins that contain all these signatures and are phylogenetically closely related to the novel-type Cld of *N. winogradskyi* exist in a large number of other microbes, most of which are denitrifiers or nitrite oxidizers. Key enzymes of these metabolisms, nitrate reductases and nitrite oxidoreductases, are known to reduce chlorate to chlorite in a side reaction. As the novel chlorite dismutases may confer resistance to the resulting chlorite, the organisms which possess these enzymes could be previously overlooked players in the bioremediation of (per)chlorate and chlorite at contaminated sites in the environment.

Keywords: Catalytic mechanism / Chlorate bioremediation / Chlorite dismutase / Crystal structure / Enzyme kinetics / Nitrite-oxidizing bacteria

Introduction

Perchlorate (ClO_4^-), chlorate (ClO_3^-), and chlorite (ClO_2^-) are a serious environmental concern as rising concentrations of these harmful compounds have been detected in groundwater, surface waters, and soils (Coates and Achenbach, 2004). While significant natural sources of perchlorate are restricted to mineral deposits in Chile, environmental contamination with this compound results from its extensive use as oxidizer in pyrotechnics and rocket fuel and its presence in certain fertilizers (Collette *et al.*, 2003). Intake of perchlorate by humans occurs mainly via drinking water, milk and certain plants and should be minimized as this chemical affects hormone production by the thyroid gland (Stanbury and Wyngaarden, 1952). Chlorate and chlorite are

used as bleaching agents in the textile, pulp and paper industries, as disinfectants and components of cleaning solutions, in pesticides, and some other applications such as chemical oxygen generators. Due to their oxidative nature, chlorate and chlorite react easily with organic material and thus have toxic effects on living cells (Ueno *et al.*, 2000). Interestingly, some microorganisms are able to use (per)chlorate as terminal electron acceptors for anaerobic respiration, leading to the reduction of perchlorate to chlorate and subsequently to chlorite (reviewed by Coates and Achenbach, 2004). These (per)chlorate-reducing bacteria (PCB) also possess the unique enzyme chlorite dismutase (Cld) that detoxifies chlorite by converting ClO_2^- to Cl^- and O_2 (van Ginkel *et al.*, 1996). In combination, (per)chlorate reduction to chlorite and the activity of Cld enable the microbially mediated removal of these compounds, making bioremediation the primary approach for the treatment of (per)chlorate contaminations (Coates and Achenbach, 2004).

Until 2008, all isolated PCB were facultatively anaerobic heterotrophs from different subclasses of the *Proteobacteria*. However, the recent description of a novel member of the genus *Moorella* (phylum *Firmicutes*) capable of (per)chlorate reduction demonstrated phylum-level diversity of PCB (Balk *et al.*, 2008). The detection of Cld activity in cell extracts from this organism indicated that this enzyme is also not confined to *Proteobacteria*. In the same year, environmental genomics and heterologous gene expression led to the identification and validation of a catalytically efficient Cld in the nitrite-oxidizing bacterium "*Candidatus* Nitrospira defluvii" (phylum *Nitrospirae*) (Maixner *et al.*, 2008). This enzyme (referred to as NdCld) is the first characterized Cld from a non-proteobacterial organism. The discovery of NdCld was unexpected, because nitrite-oxidizing bacteria (NOB) have not been considered to be relevant for (per)chlorate removal. Encompassing phylogenetic analyses made in the same study (Maixner *et al.*, 2008) showed that many of the sequenced bacterial and archaeal genomes, including those from several pathogenic bacteria, encode Cld-like proteins, although most of the respective organisms have never been observed to use (per)chlorate or convert chlorite. Most of these Cld-like proteins have not been characterized yet, but the enzyme from *Thermus thermophilus* was found to have only a very weak chlorite-degrading activity (Ebihara *et al.*, 2005). Consistently, this Cld-like protein is phylogenetically not closely related to the catalytically efficient Clds from PCB and *Ca*. N. defluvii in phylogenetic trees (Maixner *et al.*, 2008).

Recently, we determined the crystal structure of NdCld and found a high degree of structural conservation compared to Clds from *Proteobacteria* (Kostan *et al.*, in press). Cld is a heme

enzyme that, based on the published structural information, consists of five or six identical subunits (de Geus *et al.*, 2009; Ebihara *et al.*, 2005; Goblirsch *et al.*, 2010; Kostan *et al.*, in press). Besides its importance for bioengineering, Cld is extremely interesting from a biochemical perspective. Besides photosystem II and a yet uncharacterized enzyme of an anaerobic methane-oxidizing bacterium (Ettwig *et al.*, 2010), Cld is the only known enzyme which efficiently catalyzes the formation of a covalent O-O bond as its principal function. A recently proposed reaction mechanism for Cld starts with the cleavage of the Cl-O bond at the heme cofactor to form a ferryl porphyrin radical species (compound I) and hypochlorite (Lee *et al.*, 2008). Subsequently, the hypochlorite recombines with compound I to form Cl$^-$ and O$_2$.

In this study, we address the question of whether yet uncharacterized Cld-like proteins found in other microorganisms than the known PCB and *Nitrospira* might efficiently degrade chlorite. For this purpose, we chose to determine the crystal structure and to analyze biochemical features of a Cld-like protein from the nitrite oxidizer *Nitrobacter winogradskyi*, which like *Ca*. N. defluvii does not belong to the functional group of heterotrophic PCB. This protein (NwCld) was detected in the sequenced genome of this microbe (Starkenburg *et al.*, 2006) and is considerably smaller (183 amino acid residues) than all functionally validated (i.e. canonical) Clds (251-285 residues). Interestingly, NwCld represents a separate phylogenetic lineage of Cld-like proteins that are all similar in size and are encoded in the genomes of various *Alpha-*, *Beta-*, *Gamma-* and *Deltaproteobacteria* as well as *Cyanobacteria*. For none of the proteins affiliated with this lineage, a chlorite transforming capability has been shown yet. Due to these features, NwCld is a superb candidate for experimentally evaluating whether the phylogenetic and structural diversity of active Clds might be greater than previously anticipated. Our results demonstrate that *N. winogradskyi* encodes a structurally highly unusual and catalytically active Cld, suggesting a yet unrecognized diversity of organisms potentially involved in the bioremediation of (per)chlorate and chlorite, and thus could open new perspectives for future research on these processes.

Materials and methods

Cloning of NwCld

A DNA fragment containing the full-length coding region of chlorite dismutase from *Nitrobacter winogradskyi* (accession no. YP_319047) was amplified by PCR using the newly designed primers NwCldF (5'-CGA GCG CAT ATG ACG TTC ACA GTC TTC ACC-3') and NwCldR (5'-GCG CGA GGA TCC CCT ATC GCG CGC GCC AAT CG-'3) and *N. winogradskyi*

genomic DNA as template. The amplicon was cloned into expression vector pET-21b(+) (Merck/Novagen, Darmstadt, Germany) for the subsequent production of a C-terminally His-tagged fusion protein.

Heterologous expression and purification of NwCld

The recombinant NwCld was expressed in *E. coli* Tuner (DE3) cells (Merck/Novagen) grown in heme-enriched Luria-Bertani (LB) medium. Briefly, LB medium supplemented with carbenicillin (100 μg/ml) and hemin (50 μg/ml) was inoculated with a freshly prepared overnight culture (at a dilution ratio of 1:100). The culture was grown at 37°C under agitation (220 rpm) until the early stationary phase was reached (OD_{600}=1.8). For NwCld expression, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the final concentration of 0.1 mM, and temperature as well as the shaker speed were reduced to 20°C and 180 rpm, respectively. After 12 hrs, the culture was centrifuged and the resulting cell pellet was either processed immediately or frozen in liquid nitrogen and stored at -80°C. When needed, cell pellets were resuspended in 50 mM Tris, 300 mM NaCl, 2% glycerol, 20 mM imidazole (pH 8.5) supplemented with 1 mM phenylmethylsulfonylfluorid (PMSF) and 100 μM hemin. The resulting cell suspension was lysed by sonication and clarified by centrifugation. Subsequently, the supernatant was loaded onto 20 ml HisTrap FF crude (GE Healthcare, Vienna, Austria) columns. The eluted proteins were screened by SDS-PAGE, and fractions containing NwCld were pooled and applied onto a HiLoad 26/60 Superdex 200 pg column (GE Healthcare) equilibrated with buffer (50 mM Tris, 150 mM NaCl, pH 8.5). Aliquots of purified protein were concentrated to 11 mg/ml, frozen in liquid nitrogen, and stored at -80°C until further use.

Crystallization of NwCld

Initial screening for suitable crystallization conditions for NwCld was performed at 22°C with commercially available crystallization screens, using the sitting-drop vapor diffusion technique and a nanodrop-dispensing robot (Phoenix RE, Rigaku Europe, Kent, UK). First crystals of NwCld appeared within one week in 1.0 M Na_2HPO_4/NaH_2PO_4, pH 8.2 (condition no. 55 of the SaltRX crystallization screen, Hampton Research, Aliso Viejo, CA, USA). After several rounds of optimization trials using the same aforementioned set-up, single, well diffracting crystals of NwCld were obtained from 53.6 mM Na_2HPO_4 and 846.4 mM NaH_2PO_4. Prior to data collection,

crystals were transferred into crystallization solution supplemented with 30% glycerol and were flash-frozen in liquid nitrogen.

Data collection, processing, and phasing

Diffraction data were collected from a cubic-shaped crystal with approximate dimensions $300x200x120$ μm^3, at the ESRF beamline ID14-4 using 1.278 Å wavelength, to a maximum resolution of 2.1 Å. Diffraction data were integrated and scaled using the software XDS (Kabsch, 1993). Data collection statistics are summarized in Supplementary Table S1.

The phase problem was solved by single wavelength anomalous dispersion (SAD) methodology, exploiting the anomalous signal of iron. Heavy atom search, SAD phasing, solvent flattening and auto-building were carried out using AutoSol from the PHENIX software suite (Adams *et al.*, 2010).

Model building, refinement and validation of the structure

The structure of NwCld was refined using the programs phenix.refine of the PHENIX software package and TLSMD (Adams *et al.*, 2010; Painter and Merritt, 2006). Manual model building was performed in COOT (Emsley and Cowtan, 2004) while validation of the model was done using MOLPROBITY (Davis *et al.*, 2004). Final refinement statistics are summarized in Table S1. The structure has been deposited at the RCSB Protein Data Bank (accession no. XXXXX).

Structural analysis and superposition

Structure comparisons and superpositions were performed by using the SSM server (Protein Structure Comparison service SSM at EBI), the program SUPERPOSE of the CCP4-package and PyMOL (http://www.pymol.org) (CCP4, 1994; Krissinel and Henrick, 2004). The subunit interface of structures was analyzed using the Protein Interfaces, Surfaces and Assemblies Service PISA at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) (Krissinel and Henrick, 2007).

Steady-state kinetics

The chlorite-dismutating activity of NwCld was measured continuously using a Clark-type oxygen electrode (Oxygraph Plus, Hansatech Instruments, Norfolk, UK) inserted into a stirred

water bath kept at 30°C. The electrode was equilibrated to 100% air saturation by bubbling air to the reaction mixture for at least 15 min and for 0% air saturation by bubbling with N_2 for at least 15 min to derive an offset and calibration factor. Reactions were carried out in O_2-free 50 mM phosphate buffer (pH 7.0) with 50 μM to 100 mM $NaClO_2^-$ added from a stock made in the same buffer. Reactions were started by addition of 20 nM solutions of NwCld. With increasing chlorite concentrations, irreversible inactivation of NwCld occurred as was evident by inspection of individual time traces. Thus, it was important to (i) use only the initial linear phase for rate calculation and to (ii) deduce Michaelis-Menten parameters from a set of chlorite concentrations below 1 mM. Molecular oxygen production rates (μM O_2 s^{-1}) were obtained from initial linear time traces (<10% substrate consumed) and plotted against chlorite concentrations.

Spectra analysis and heme determination

UV-visible spectra were recorded on a Nanodrop 2000c spectrophotometer (Thermo Scientific, Waltham, MA, USA) at 22°C. Heme type and NwCld:heme *b* stoichiometry were determined by the pyridine hemochrome assay(Smith, 1975). In detail, oxidized protein samples in 50 mM potassium phosphate (pH 7.0) were mixed with pyridine and NaOH (final concentrations: 20% (v/v) pyridine and 0.1 M NaOH). After recording the oxidized spectrum, sodium dithionite was added to a final concentration of 10 mM and the spectrum of the reduced pyridine hemochrome was recorded and the difference spectrum was calculated. The heme content was calculated using the molar extinction coefficient ε_{418}=191 500 M^{-1} cm^{-1} (Smith, 1975).

Phylogenetic analyses

All phylogenetic analyses were performed by using the ARB software package (Ludwig *et al.*, 2004) and an already existing database of aligned amino acid sequences of Clds and Cld-like proteins (Maixner *et al.*, 2008). The alignment was manually refined based on the crystal structures of NwCld (this study); NdCld (Kostan *et al.*, in press); Cld of *Azospira oryzae* (AoCld, PDB: 2VXH); Cld of *Dechloromonas aromatica* (DaCld, PDB: 3M2Q and 3M2S); a Cld-like protein from *Geobacillus stearothermophilus* (GsCld, PDB: 1T0T); a Cld-like protein from *Thermoplasma acidophylum* (TaCld, PDB: 3DTZ); and a Cld-like protein from *Thermus thermophilus* (TtCld, PDB: 1VDH). Highly conserved secondary and tertiary structure motifs were used to identify homologous residues. Phylogenetic trees were calculated by applying protein maximum-likelihood [PhyML (Guindon and Gascuel, 2003)] and maximum-parsimony

(PHYLIP version 3.66 with 100 bootstrap iterations) methods, both with the JTT substitution model. In total 243 alignment columns were used for phylogenetic analysis. To determine the coverage of *cld* genes by *cld*-targeted PCR primer sets, an ARB PT-Server was established from the nucleotide gene sequences of the Clds and Cld-like proteins in the database. Primer sequences were then matched to this PT-Server database by using the ARB probe match tool. All possible sequence permutations of degenerate primers were resolved and matched by using the respective tools of ARB.

Miscellaneous methods

Protein concentrations were determined by UV_{280nm} absorption. The extinction coefficient was calculated with ProtParam (Gasteiger *et al.*, 2005). The oligomeric state in solution was determined with analytical size exclusion chromatography using a Superdex 75 10/300 (GE Healthcare) column equilibrated with 100 mM Tris buffer at pH 8.5 and 200 mM NaCl. All molecular graphics figures were prepared using PyMOL.

Results and discussion

The heterologous over-expression of NwCld in *E. coli*, and subsequent purification, yielded sufficient amounts of pure protein for successful crystallization, structure analysis, and kinetic characterization. The structure of NwCld was determined to a resolution of 2.1 Å. The protein crystallizes in the $P4_1$ space group with unit cell dimensions: a = b = 102.62 Å, c = 49.12 Å. The structural analysis revealed several interesting differences at the tertiary and quaternary structure levels compared to previously determined structures of Clds and Cld-like proteins. Most remarkably, NwCld was found to be a homodimer (Figure 1a). This is consistent with analytical size exclusion chromatography, where NwCld eluted in a single peak with an apparent molecular weight of 43 kDa that corresponds to an NwCld dimer (data not shown). In contrast, all canonical Clds and all Cld-like proteins, whose structures have been determined, are homopentamers or homohexamers, respectively (de Geus *et al.*, 2009; Ebihara *et al.*, 2005; Goblirsch *et al.*, 2010; Kostan *et al.*, in press). In addition, the NwCld primary sequence is about 30% shorter compared to the canonical Clds, with a significant deletion in the N-terminal region. These unique structural properties of NwCld are also reflected by the phylogenetic analysis, which revealed that NwCld does not cluster with any other structurally characterized member of the Cld-like protein

superfamily (Figure 2). Due to weak electron density, we could not model residues 41-47 in monomer A and residues 41-48 in monomer B, suggesting that these regions are flexible and were disordered in the crystal lattice. Superposing NwCld and all other available structures of Clds and Cld-like proteins revealed a high structural similarity at the monomer level, with a root mean square deviation (r.m.s.d) of 1.40 Å over 119 superposed Cα atoms, although the amino acid sequence identities between NwCld and these proteins are as low as 13-27%. Consistent with the phylogeny of Cld and Cld-like proteins (Figure 2), NwCld is structurally most similar to DaCld from *D. aromatica* (structure 3M2Q: r.m.s.d of 1.70 Å and 3M2S: r.m.s.d of 1.70 Å) and to NdCld (r.m.s.d of 1.73 Å) and least similar to TaCld from *T. acidophilum* (r.m.s.d of 2.25 Å).

Monomer structure of NwCld

All previously determined subunit structures of Clds and Cld-like proteins (de Geus *et al.*, 2009; Ebihara *et al.*, 2005; Goblirsch *et al.*, 2010; Kostan *et al.*, in press) consist of two similar domains with a ferredoxin like fold. Both domains are characterized by a four-stranded antiparallel β-sheet flanked by six α-helices. The two β-sheets from each domain pack together at an angle of about 65°, forming a central flattened β-barrel surrounded on both sides by α-helices (Figure 1b). NwCld, however, lacks the major part of the N-terminal domain (Figure 1c). In particular, all N-terminal α-helices are missing and only the β-sheet has been conserved. Furthermore, also this β-sheet is different: the β2-strand, which in canonical Clds interacts with the β3-strand, is no longer part of the first β-sheet but rather forms the fifth strand of the second β-sheet interacting with the β4'-strand of the C-terminal ferredoxin-like domain (Figure 1c). Therefore, the central β-barrel no longer consists of two similar four-stranded β-sheets but of one three-stranded and one five-stranded β-sheet. The C-terminal domain is structurally highly similar to the C-terminal domains of previously described Clds (Figure 1b, c) although it lacks the last α-helix α6' that is present in the canonical Clds (Figure 1b-d).

Active site structure of NwCld

Similar to other known Cld structures, the active site of NwCld with the bound heme *b* is located in the C-terminal domain of the molecule, in a cavity formed between the second β-sheet (β1' to β4') and helices α2', α3'and α4' (Figure 1c). Most residues within 4 Å from the heme *b* have a hydrophobic nature and interact with the prosthetic group *via* van der Waals contacts.

Furthermore, several residues are involved in specific interactions that are important for heme binding and enzyme function. On the proximal side of the heme *b*, His114 from the α3' helix coordinates the heme iron at a distance of 2.16 Å (Figure 1e) as well as forms a hydrogen bond with Glu167. The latter residue is hydrogen bonded to Lys92, which is also within hydrogen bonding distance to one propionate group of the heme *b* (Figure 1e). This intricate hydrogen bonding network increases the imidazolate character of the proximal histidine residue, thereby shifting the reduction potential of the heme iron to more negative values. This might be important for the stabilization of higher heme oxidation state(s) involved in chlorite dismutation (Lee *et al.*, 2008). The second propionate of heme *b* is hydrogen bonded to the backbone of two residues located in the loop between β4 and α1': Arg60 and Tyr61, of which Arg60 is also important for the stability of the dimer interface (see also Supplementary information and Supplementary Figure S4b). On the proximal heme side at a distance of 3.8 Å from one of the propionate groups there is a pair of tryptophan residues (Trp96 and Trp97), whose aromatic ring systems are oriented perpendicularly to each other (Figure 1e). Trp96 may play a role in the catalytic mechanism, acting as the electron donor for the reduction of compound I to compound II (Lee *et al.*, 2008). The second tryptophan (Trp97) contributes together with Lys92, Met 99, Arg104 and Leu107 to the stabilization of the Trp96 side chain orientation that is suitable for the interaction with the heme *b* propionate group. In addition, Trp97 influences the electronic structure of Trp96 through interaction of its partially positively charged hydrogen atoms with π electrons and partially negatively charged carbon atoms of Trp96.

At the distal side, the heme iron is coordinated by a water molecule (Wat108) at distance of 2.9 Å. This water molecule is linked *via* a hydrogen bonding network to a second water molecule (Wat95) and Arg127 (Figure 1e). The Arg127 side chain, which points away from the heme, is further stabilized by hydrogen bonds to four additional water molecules (Wat1, Wat174, Wat193 and Wat260) and to the side chain of Gln74 (Supplementary Figure S1a). In our previous study on NdCld, an arginine residue at this position was experimentally shown to be important for the catalytic activity, most likely by regulating substrate uptake and by stabilizing a reaction intermediate (Kostan *et al.*, in press). Consistent with its functional importance, this arginine residue occurs in all catalytically efficient Clds (see also below) and its side chain conformation is stabilized by a similar pattern of interactions in these enzymes (see Supplementary information). At the inner part of the active site cavity the two vinyl groups of the heme ligand

interact with Ile88 and Leu122, respectively. These two amino acid residues define the depth of the ligand cavity and are therefore important for heme binding (Figure 1e).

Additional structural details of the active site and the dimer interface of NwCld are provided as Supplementary information (Supplementary Text and Supplementary Figures S1, S3 and S4).

Heme analysis and UV-Vis spectral characterization

Absorption spectra of reduced pyridine hemochromes of NwCld showed maxima at 418 nm (Soret), 526 nm (β) and 557 (α), typical for protoporphyrin IX (heme *b*) (Smith, 1975) as was seen in other so far characterized Clds (de Geus *et al.*, 2009; Kostan *et al.*, in press). The heme content was determined to be 0.58 heme *b* per NwCld monomer the Reinheitszahl A_{407nm}/A_{280nm} = 1.9). This ratio is lower than expected since the crystal structure is fully loaded with heme. Incomplete heme occupancy in recombinant proteins has been reported for other Clds (Mehboob *et al.*, 2009) and might be related with bottlenecks in heme supply by the host due to high protein overexpression and/or suboptimal cultivation conditions. Native NwCld showed a Soret maximum at 407 nm which shifted to 435 nm upon reduction by dithionite (Supplementary Figure S2). Similar Soret maxima of ferric and ferrous proteins have been reported for other Clds (Maixner *et al.*, 2008; Stenklo *et al.*, 2001; van Ginkel *et al.*, 1996).

Steady-state kinetics of NwCld

The stoichiometry of the Cld reaction has been reported to be 1 mol Cl⁻ and 1 mol O_2 out of 1 mol ClO_2^- (Lee *et al.*, 2008). Figure 3a depicts the plot of the initial rate of O_2 release by NwCld *versus* chlorite concentration. It demonstrates saturation in the initial rate (v_0) with increasing chlorite concentration followed by a decrease in v_0 at very high chlorite concentrations (> 100 mM), which indicates uncompetitive substrate inhibition. At these high concentrations, where the inhibitory effect predominates, an inhibition constant (K_i = 310 mM) could be determined from the Dixon plot depicted in Figure 3e. Inactivation of NwCld by chlorite was evident by inspection of the individual time traces and the total amount of released dioxygen at defined $[ClO_2^-]$. Only at low chlorite concentrations (<50 μM) the substrate was completely consumed, whereas at higher concentrations residual substrate remained due to inactivation of the enzyme in the course of the reaction (Figure 3b). At $[ClO_2^-]$ > 1 mM less than 10% of chlorite was converted to Cl⁻ and O_2. Thus, in order to minimize interference with inhibition kinetics, Michaelis-Menten parameters were calculated from a set of chlorite concentrations with $[ClO_2^-]$ ≤

1 mM, which probably reflect physiological conditions better than higher concentrations (for better clarity see also the semilogarithmic plot in Figure 3c). Michaelis-Menten parameters were determined from the double-reciprocal plot (Figure 3d), and K_M and k_{cat} were calculated to be 90 µM and 190 s^{-1}, resulting in a catalytic efficiency (k_{cat}/K_M) of 2.1 × 10^7 M^{-1} s^{-1}. These values resemble the kinetic parameters of the canonical Clds. Enzymes from PCB, and NdCld, have a K_M of 58 (NdCld) or 80 to 260 µM (PCB), a k_{cat} of 35 (NdCld) or 230 to 1 880 s^{-1} (PCB), and a catalytic efficiency of 6 × 10^5 (NdCld) or 2.7 to 35.4 × 10^6 M^{-1} s^{-1} (PCB) (Kostan *et al.*, in press; Mehboob *et al.*, 2009; Stenklo *et al.*, 2001; Streit and DuBois, 2008). For comparison, the kinetic parameters of the Cld-like protein TtCld from *T. thermophilus* with a weak chlorite-dismutating activity are 13 mM, 0.77 s^{-1}, and 59 M^{-1} s^{-1} (Ebihara *et al.*, 2005). Thus, despite the pronounced structural differences to the canonical Clds (Figure 1), NwCld is an efficient chlorite dismutase. This result is consistent with the phylogenetic analysis. The distinct and well-supported lineage that contains NwCld shares a common ancestor with another lineage containing all known and highly efficient canonical Clds to the exclusion of all other Cld-like proteins (Figure 2). Here we refer to the canonical Clds as "lineage I", whereas the cluster formed by NwCld and similar proteins is designated "lineage II" (Figure 2). The monophyly of lineages I and II indicates that both groups share steps of their evolutionary history, which most likely were key to the development of efficient chlorite-dismutating enzymes and separated these lineages from the other Cld-like proteins. As the small size of NwCld and the other lineage II proteins is an exception within the Cld-like protein superfamily (Supplementary Figure S5), it does probably not represent the ancestral state but was rather caused by a partial loss of the N-terminal domain during the evolution of lineage II.

The temperature optimum of the chlorite dismutating activity of NwCld was found to be at 20°C, which is lower than the temperature optima of NdCld [25°C; (Maixner *et al.*, 2008)] and Clds from PCB [25-30°C; (Mehboob *et al.*, 2009; van Ginkel *et al.*, 1996)]. The pH optimum of NwCld (pH 5.5) is slightly lower than that of other Clds [pH 6.0; (Mehboob *et al.*, 2009; van Ginkel *et al.*, 1996)].

Signature residues of catalytically efficient Clds
The currently available structures and kinetic data for Clds and Cld-like proteins allowed us to identify potential signature residues that could define a catalytically efficient Cld. Some of the

aforementioned residues in the vicinity of the active site of NwCld (Lys92, Trp96, and His114; NwCld numbering) are strictly conserved in all Clds and Cld-like proteins (Supplementary Figure S5). However, the five residues Ile88, Trp97, Leu122, Arg127, and Glu167 are conserved only in NwCld, the canonical Clds (including NdCld), and in sequences that are closely related to these reference proteins (Supplementary Figure S5). As outlined above, these residues are proposed to stabilize the heme cofactor or to play roles in the catalytic mechanism of Cld. For Arg127 an important function for catalysis has been experimentally demonstrated (Kostan *et al.*, in press). All Cld-like proteins whose structure has been determined and found to be without heme, or which have been tested for Cld activity but found to be inactive (TtCld, GsCld, and TaCld), have different residues at these key positions (Supplementary Figure S5). In these proteins Arg127 is replaced by a polar residue (glutamine or serine). Glu167, which stabilizes the proximal His114 and connects it *via* a network of hydrogen bonds to Lys92 (see above), is replaced by a hydrophobic residue (alanine or valine) that cannot form a similar network. At the position corresponding to Ile88, these Cld-like proteins have a tyrosine residue. *In silico* mutation of Ile88 to tyrosine leads to a steric clash between the tyrosine side chain and the vinyl group of the heme moiety. Similarly, the second vinyl group in NwCld interacts with Leu122, while the Cld-like protein structures without heme (GsCld, TaCld and TtCld) have an alanine, a proline or a glutamine in the corresponding position, leading to less favorable interactions with a putative heme *b*. Moreover, in these Cld-like proteins the loop between strand β4 and helix α1', whose backbone interacts with a heme propionate group in NwCld (see above), is in a different position (not shown). These subtle differences suggest that if these Cld-like proteins bind heme *b*, they might bind it weaker and in a slightly different orientation than in the canonical Clds and NwCld. Assuming these proteins still bind heme, it would be expected that their Cld activity is low due to the differences in the key residues discussed above.

Given the high diversity of Cld-like proteins [Figure 2 and Maixner *et al.* (2008)] and the relatively low degree of raw sequence conservation in this superfamily, the aforementioned five signature residues are valuable indicators for a tentative functional classification of novel Cld-like sequences. For example, the presence of these signatures indicates that not only NwCld, but also all known lineage II-proteins from other bacteria (Figure 2, Supplementary Figure S5) could be efficient Clds. Interestingly, this group includes other nitrite oxidizers (*Nitrobacter* strain Nb-311A and *Nitrococcus mobilis*) and important pathogens of humans (*Klebsiella pneumoniae* and *Pseudomonas aeruginosa*) (Figure 2). However, future experimental work is needed to validate

the activity of additional lineage II-Clds and to verify the proposed functional importance of all conserved signature residues by site-directed mutagenesis (Kostan *et al.*, in press).

Ecophysiological function of NwCld and related proteins

The validation of NwCld as a chlorite dismutase, and the finding that the entire lineage II likely represents a novel class of Clds, raise the question of the ecophysiological importance of these enzymes. Remarkably, no organism carrying a lineage II Cld (Figure 2) is known to grow with (per)chlorate as terminal electron acceptor and we cannot exclude the possibility that some of these enzymes have yet unrecognized biological functions unrelated to chlorite dismutation. However, most of the bacteria with lineage II Cld can reduce nitrate using membrane-bound (Nar) and/or periplasmic (Nap) dissimilatory nitrate reductases (Figure 2). In many nitrate-reducing bacteria, chlorate is transformed to chlorite by nitrate reductase due to the structural similarity between chlorate and nitrate [Logan *et al.* (2001) and references cited therein]. It is generally assumed that chlorate reduction is confined to Nar-possessing nitrate reducers and that the Nap enzyme does not reduce chlorate. However, this distinction might not hold always true as in *Paracoccus denitrificans* cells expressing only Nap and not Nar, effective inhibition by chlorate of nitrate reduction was observed (Kucera, 2006) and a periplasmic nitrate reductase of *Rhodobacter sphaeroides* exhibits chlorate reductase activity (Castillo *et al.*, 1996). Chlorate reduction does not yield energy in most nitrate reducers, but the resulting chlorite production is responsible for the toxicity of chlorate to those bacteria that have nitrate reductase and no Cld (Coates and Achenbach, 2004). Thus, it is tempting to speculate that lineage II-Clds of nitrate-reducing (e.g., denitrifying) bacteria, which grow at chlorate-contaminated sites, may offer protection from chlorite produced due to chlorate reduction by their nitrate reductases. Such a protective function of Cld could also be beneficial to *Nitrobacter* and *Nitrococcus*, because their key enzyme nitrite oxidoreductase (Nxr) (Tanaka *et al.*, 1983) is highly similar to Nar (Kirstein and Bock, 1993) and does also reduce chlorate (Meincke *et al.*, 1992). Interestingly, previous studies found that in absence of O_2, *Nitrobacter winogradskyi* and an environmental *Nitrobacter* isolate oxidized nitrite using chlorate as electron acceptor (Hynes and Knowles, 1983; Lees and Simpson, 1957). However, this activity ceased after a few hours and nitrite oxidation became inhibited by chlorite. Apparently, these results contradict our finding that *N. winogradskyi* has an efficient Cld, which should confer some resistance to chlorite. On the other hand, the aforementioned experiments with *Nitrobacter* were performed in nitrite media containing

relatively high chlorate concentrations [4.2-17 mM (Lees and Simpson, 1957) or 10 mM (Hynes and Knowles, 1983)]. Similarly, nutrient media used to isolate PCB have been prepared with 10 mM chlorate (e.g., Achenbach *et al.*, 2001), selecting for specialized organisms adapted to grow with such high chlorate concentrations and to detoxify the resulting amounts of chlorite. However, nitrate-reducing bacteria or nitrite oxidizers may only be protected by their Cld from lower chlorite concentrations. Consistently, *Nitrobacter* is strongly inhibited by chlorite concentrations above 1 mM but inhibition was only weak with 0.3 mM (Lees and Simpson, 1957) or less than 0.1 mM (Hynes and Knowles, 1983). In this context, it is noteworthy that toxic effects of chlorite on other bacteria have been observed already at concentrations of 0.01-0.02 mM (van Wijk *et al.*, 1998). As chlorate and nitrate are competing substrates of nitrate-reducing enzymes, the ratio of their concentrations should also be an important factor determining how much chlorite is produced. Hence, a protective effect of Cld in non-PCB might be masked in laboratory experiments with high chlorate and low nitrate concentrations in the applied media.

Assuming a protective function of Clds in certain nitrite oxidizers and nitrate reducers, we hypothesize that microbial communities, which carry out (per)chlorate bioremediation, are more complex than previously known. This would be a consequence of the diversity (Figure 2) and the possibly high abundance of organisms carrying nitrate reductases or nitrite oxidoreductases in combination with putative lineage II *cld* genes. The latter genes were found not only in sequenced genomes from pure cultures, but also on some metagenomic reads from marine systems and a freshwater sediment (Figure 2) suggesting that they originate from highly abundant organisms in the respective samples. Moreover, one lineage II *cld* gene was found on plasmid pAKD4 (Sen *et al.*, 2010) (Figure 2). This plasmid originates from soil (Dronen *et al.*, 1998) and belongs to incompatibility group IncP-1, which comprises highly promiscuous broad-host-range plasmids that replicate in different classes of the *Proteobacteria* (Sen *et al.*, 2010). Thus, lineage II *cld* genes likely are subject to horizontal transfer on highly mobile genetic elements, which could easily promote their distribution among phylogenetically diverse bacteria in the environment. However, the relative contribution of organisms carrying such genes to chlorate reduction and chlorite removal will probably vary and will strongly depend on several biological and environmental parameters. These may include the species- or strain-specific tolerance against chlorate and chlorite, ambient concentrations of these substances and of nitrate, enzymatic traits of Clds and nitrate reductases, and available electron donors and carbon sources. Under dynamic

environmental conditions, shifts in some of these parameters could lead to a complex temporal and spatial heterogeneity of the population structure and metabolic activity in (per)chlorate-removing communities. Thus, understanding the microbiology of (per)chlorate bioremediation will require reliable tools to detect, identify, and monitor such potentially chlorate- and chlorite-removing organisms in addition to the known PCB. The *cld* gene has been used as a functional marker for the cultivation-independent detection of PCB in environmental samples by using specific PCR primers (Bender *et al.*, 2004). The design of all primers was based on the sequences of lineage I *cld* genes from proteobacterial PCB. Considering also new primers specific for the *cld* of *Ca*. N. defluvii (Maixner *et al.*, 2008), lineage I is well covered by the available primer sets (Table S2). However, all lineage II *cld* genes have numerous base mismatches to all published primers (data not shown) and thus were overlooked when PCR was applied to detect *cld* genes in environmental samples or laboratory cultures. The case of *Pseudomonas stutzeri* and *P. chloritidismutans* demonstrates this problem. The only phenotypic trait that differentiates these two closely related pseudomonads is the ability of *P. chloritidismutans* to use chlorate but not nitrate as electron acceptor, whereas the denitrifier *P. stutzeri* grows with nitrate but not with chlorate (Cladera *et al.*, 2006). *P. chloritidismutans* possesses a lineage I Cld (Figure 2), and the corresponding gene can be amplified (Cladera *et al.*, 2006) by using lineage I-specific primers (Bender *et al.*, 2004). These primers did not yield any PCR amplicon with 27 *P. stutzeri* strains (Bender *et al.*, 2004; Cladera *et al.*, 2006), but the only sequenced genome of *P. stutzeri* from strain A1501 contains a lineage II putative *cld* gene (Figure 2). Strain A1501 was not included in the aforementioned PCR screenings, but it is not unlikely that several of the tested strains also contained a lineage II *cld* gene, and as nitrate-reducing denitrifiers with a functional Cld they might contribute to chlorate and chlorite removal *in situ*. This potential trait would also be overlooked by cultivation-based approaches, because *P. stutzeri* does not grow with chlorate as electron acceptor (see above).

Conclusions

Recent progress in genome and metagenome sequencing has enormously increased the size and phylogenetic complexity of many groups of homologous proteins including the Cld-like protein superfamily (Maixner *et al.*, 2008). However, for many of these groups it remains unclear whether all affiliated proteins share a specific function or whether modified functions occur in different species or evolutionary lineages. Tackling this question is often extremely difficult, in

particular because most microorganisms cannot be grown under laboratory conditions and/or no genetic tools are available for their manipulation. For the Cld-like protein superfamily only the canonical Clds of lineage I had been assigned a biological function in PCB, whereas the substrates and activities of the numerous other Cld-like enzymes remained an enigma. In this study, by combining heterologous expression with structural and biochemical analyses, we identified the *N. winogradskyi* Cld-like protein representing another lineage from this superfamily as a novel and structurally distinct chlorite dismutase. This result is interesting from the biochemical perspective, but also has implications for our understanding of the microbiology of (per)chlorate and chlorite bioremediation. We do not know at present whether organisms, which possess a lineage II-Cld, make a significant contribution to these processes. However, the relatively high efficiency of NwCld, together with the detected presence of similar genes in diverse organisms and on a mobile genetic element, implies that this possibility should not be overlooked in future research. The design of such studies should thus be broader than that of previous approaches to investigate PCB. Instead of exclusively isolating the most robust and fast-growing "key players", another goal will be to detect the possibly subtle activities of phylogenetically diverse and physiologically versatile organisms, which are less tolerant to chlorate and chlorite. Such microbes may be relevant mainly at the interfaces of pristine and contaminated areas or after most of the pollutants have been removed by "professional" PCB. As the combined numbers of these organisms, for example in soil, could result in quite large amounts their contribution may be important even if chlorate and chlorite transformations are merely side reactions in their metabolism.

Supplementary information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej).

References

Achenbach LA, Michaelidou U, Bruce RA, Fryman J, Coates JD (2001). *Dechloromonas agitata* gen. nov., sp. nov. and *Dechlorosoma suillum* gen. nov., sp. nov., two novel environmentally dominant (per)chlorate-reducing bacteria and their phylogenetic position. *Int. J. Syst. Evol. Microbiol.* **51:** 527-533.

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N *et al* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66:** 213-221.

Balk M, van Gelder T, Weelink SA, Stams AJM (2008). (Per)chlorate reduction by the thermophilic bacterium, *Moorella perchloratireducens* sp. nov., isolated from an underground gas storage. *Appl. Environ. Microbiol.* **74:** 403-409.

Bender KS, Rice MR, Fugate WH, Coates JD, Achenbach LA (2004). Metabolic primers for detection of (Per)chlorate-reducing bacteria in the environment and phylogenetic analysis of *cld* gene sequences. *Appl. Environ. Microbiol.* **70:** 5651-5658.

Castillo F, Dobao MM, Reyes F, Blasco R, Roldan MD, Gavira M *et al* (1996). Molecular and regulatory properties of the nitrate reducing systems of *Rhodobacter*. *Curr. Microbiol.* **33:** 341-346.

CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50:** 760-763.

Cladera AM, Garcia-Valdes E, Lalucat J (2006). Genotype versus phenotype in the circumscription of bacterial species: the case of *Pseudomonas stutzeri* and *Pseudomonas chloritidismutans*. *Arch. Microbiol.* **184:** 353-361.

Coates JD, Achenbach LA (2004). Microbial perchlorate reduction: rocket-fueled metabolism. *Nat. Rev. Microbiol.* **2:** 569-580.

Collette TW, Williams TL, Urbansky ET, Magnuson ML, Hebert GN, Strauss SH (2003). Analysis of hydroponic fertilizer matrixes for perchlorate: comparison of analytical techniques. *Analyst* **128:** 88-97.

Davis IW, Murray LW, Richardson JS, Richardson DC (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32:** W615-619.

de Geus DC, Thomassen EA, Hagedoorn PL, Pannu NS, van Duijn E, Abrahams JP (2009). Crystal structure of chlorite dismutase, a detoxifying enzyme producing molecular oxygen. *J. Mol. Biol.* **387:** 192-206.

Dronen AK, Torsvik V, Goksoyr J, Top EM (1998). Effect of mercury addition on plasmid incidence and gene mobilizing capacity in bulk soil. *FEMS Microbiol. Ecol.* **27:** 381-394.

Ebihara A, Okamoto A, Kousumi Y, Yamamoto H, Masui R, Ueyama N *et al* (2005). Structure-based functional identification of a novel heme-binding protein from *Thermus thermophilus* HB8. *J. Struct. Funct. Genomics* **6:** 21-32.

Emsley P, Cowtan K (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60:** 2126-2132.

Ettwig KF, Butler MK, Le Paslier D, Pelletier E, Mangenot S, Kuypers MM *et al* (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* **464:** 543-548.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD *et al* (2005). Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed). *The proteomics protocols handbook*. Humana Press. pp 571-607.

Goblirsch BR, Streit BR, Dubois JL, Wilmot CM (2010). Structural features promoting dioxygen production by *Dechloromonas aromatica* chlorite dismutase. *J. Biol. Inorg. Chem.*

Guindon S, Gascuel O (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52:** 696-704.

Hynes RK, Knowles R (1983). Inhibition of chemoautotrophic nitrification by sodium chlorate and sodium chlorite: a reexamination. *Appl. Environ. Microbiol.* **45:** 1178-1182.

Kabsch W (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26:** 795-800.

Kirstein K, Bock E (1993). Close genetic relationship between *Nitrobacter hamburgensis* nitrite oxidoreductase and *Escherichia coli* nitrate reductases. *Arch. Microbiol.* **160:** 447-453.

Kostan J, Sjöblom B, Maixner F, Mlynek G, Furtmüller PG, Obinger C *et al* (in press). Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium "*Candidatus* Nitrospira defluvii": Identification of a catalytically important amino acid residue. *J. Struct. Biol.*

Krissinel E, Henrick K (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60:** 2256-2268.

Krissinel E, Henrick K (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372:** 774-797.

Kucera I (2006). Interference of chlorate and chlorite with nitrate reduction in resting cells of *Paracoccus denitrificans.* *Microbiology* **152:** 3529-3534.

Lee AQ, Streit BR, Zdilla MJ, Abu-Omar MM, DuBois JL (2008). Mechanism of and exquisite selectivity for O-O bond formation by the heme-dependent chlorite dismutase. *Proc. Natl. Acad. Sci. U. S. A.* **105:** 15654-15659.

Lees H, Simpson JR (1957). The biochemistry of the nitrifying organisms. V. Nitrite oxidation by *Nitrobacter.* *Biochem. J.* **65:** 297-305.

Logan BE, Zhang H, Mulvaney P, Milner MG, Head IM, Unz RF (2001). Kinetics of perchlorate- and chlorate-respiring bacteria. *Appl. Environ. Microbiol.* **67:** 2499-2506.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhu K *et al* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* **32:** 1363-1371.

Maixner F, Wagner M, Lücker S, Pelletier E, Schmitz-Esser S, Hace K *et al* (2008). Environmental genomics reveals a functional chlorite dismutase in the nitrite-oxidizing bacterium *'Candidatus* Nitrospira defluvii'. *Environ. Microbiol.* **10:** 3043-3056.

Mehboob F, Wolterink AF, Vermeulen AJ, Jiang B, Hagedoorn PL, Stams AJ *et al* (2009). Purification and characterization of a chlorite dismutase from *Pseudomonas chloritidismutans*. *FEMS Microbiol. Lett.* **293:** 115-121.

Meincke M, Bock E, Kastrau D, Kroneck PMH (1992). Nitrite oxidoreductase from *Nitrobacter hamburgensis*: redox centers and their catalytic role. *Arch. Microbiol.* **158:** 127-131.

Painter J, Merritt EA (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* **62:** 439-450.

Sen D, Yano H, Suzuki H, Krol JE, Rogers L, Brown CJ *et al* (2010). Comparative genomics of pAKD4, the prototype IncP-1delta plasmid with a complete backbone. *Plasmid* **63:** 98-107.

Smith KM (1975). *Porphyrins and metalloporphyrins*. Elsevier Scientific: Amsterdam.

Stanbury JB, Wyngaarden JB (1952). Effect of perchlorate on the human thyroid gland. *Metabolism* **1:** 533-539.

Starkenburg SR, Chain PS, Sayavedra-Soto LA, Hauser L, Land ML, Larimer FW *et al* (2006). Genome sequence of the chemolithoautotrophic nitrite-oxidizing bacterium *Nitrobacter winogradskyi* Nb-255. *Appl. Environ. Microbiol.* **72:** 2050-2063.

Stenklo K, Thorell HD, Bergius H, Aasa R, Nilsson T (2001). Chlorite dismutase from *Ideonella dechloratans*. *J. Biol. Inorg. Chem.* **6:** 601-607.

Streit BR, DuBois JL (2008). Chemical and steady-state kinetic analyses of a heterologously expressed heme dependent chlorite dismutase. *Biochemistry* **47:** 5271-5280.

Tanaka Y, Fukumori Y, Yakamaka T (1983). Purification of cytochrome a_1c_1 from *Nitrobacter agilis* and characterization of nitrite oxidation system of the bacterium. *Arch. Microbiol.* **135:** 265-271.

Ueno H, Oishi K, Sayato Y, Nakamuro K (2000). Oxidative cell damage in Kat-sod assay of oxyhalides as inorganic disinfection by-products and their occurrence by ozonation. *Arch. Environ. Contam. Toxicol.* **38:** 1-6.

van Ginkel CG, Rikken GB, Kroon AG, Kengen SW (1996). Purification and characterization of chlorite dismutase: a novel oxygen-generating enzyme. *Arch. Microbiol.* **166:** 321-326.

van Wijk DJ, Kroon SG, Garttener-Arends IC (1998). Toxicity of chlorate and chlorite to selected species of algae, bacteria, and fungi. *Ecotoxicol. Environ. Saf.* **40:** 206-211.

Titles and legends to figures

Figure 1 Structural analysis of NwCld. (**a**) Crystal structure of the NwCld holoenzyme. The ribbon representation of the structure is viewed perpendicular to the vertical 2-fold symmetry axis. Monomers are shown in different colors. The heme group is shown as an orange sticks model in either monomer. The iron is displayed as an orange sphere. The N- and C-termini are labeled. Parts of the structure between Ser40 and Thr48 in one subunit and Ser40 and Pro49 in the other subunit could not be modeled in the structure due to weak electron density in these regions. (**b**) Structure of one monomer of NdCld from *Ca*. N. defluvii. The N- and C-termini are labeled, the heme *b* is presented as orange sticks with iron shown as an orange sphere. Secondary structure elements are labeled according to Kostan *et al*. (in press). (**c**) Structure of one NwCld monomer. The N-terminal domain is shown in red, whereas the C-terminal domain is shown in cyan. The labeling of secondary structure elements follows the labeling used for NdCld as shown in (b). Heme group and iron are presented as in (b). (**d**) A monomer of NdCld (shown in green and grey) superimposed onto a monomer of NwCld (shown in red and cyan). The orientation of the monomers is the same as in (b) and (c). Structural elements missing in the monomer structure of NwCld compared to the NdCld monomer structure are depicted in grey. (**e**) Stereo view of the active site of NwCld. The figure shows selected residues involved in heme *b* binding and in the catalytic mechanism of chlorite dismutation. Carbon, oxygen, and nitrogen atoms are depicted in green, red, and blue, respectively. The heme iron and water molecules (*W*) are shown as red spheres. The hydrogen bonding network spanning from Arg127 to Glu167 is visualized by dashed lines.

Figure 2 Maximum likelihood phylogenetic tree based on the amino acid sequences of selected Clds and Cld-like proteins. Names printed in boldface represent proteins with known crystal structures. Validated and catalytically efficient Clds are marked with an asterisk. The proposed phylogenetic lineages I (canonical Clds, mainly from PCB) and II (NwCld and related proteins) are delimited by curly brackets. Black circles on tree nodes symbolize high parsimony bootstrap support ($\geq 90\%$) based on 100 iterations. Accession numbers are indicated for all sequences. For lineage II, the presence of genes homologous to dissimilatory membrane-bound nitrate reductase (Nar), periplasmic nitrate reductase (Nap), or nitrite oxidoreductase (Nxr) in the respective

organisms is indicated. [1]Uncertain due to limited sequence homology to reference protein sequences. [2]These genomes contain only homologs to assimilatory nitrate reductases.

Figure 3 Kinetics of chlorite dismutation activity by NwCld. (**a**) Plot of the initial rate (v_0) of molecular oxygen evolution as a function of chlorite concentration. The inset depicts the corresponding semi-logarithmic plot. Points represent averages of three measurements. In addition, double-rectangular hyperbolar fits are shown in grey. Applied conditions were: 50 mM phosphate buffer pH 7.0, 20 nM NwCld, 30°C. (**b**) Selected time traces at different chlorite concentrations. Conditions as in (a). (**c**) Double reciprocal plot of v_0 *versus* chlorite concentration for the determination of kinetic parameters. (**d**) Dixon plot for the determination of K_I.

Supplementary Figure S1 Environment of the catalytically important arginine residue as found in the structures of (**a**) dimeric NwCld, (**b**) pentameric NdCld, and (**c**) hexameric AoCld. The NwCld structure represents the enzyme in its native state, as a water molecule (Wat108) was found to coordinate with the heme iron (panel a). Instead of water, imidazole (IMD) and thiocyanate (SCN) are bound to the heme iron of NdCld (panel b) and AoCld (panel c), respectively. In the case of NdCld, water molecules that stabilize the arginine residue in the native enzyme were replaced by ethylene glycol (EDO) from the cryo solution. Carbon atoms are depicted in green, cyan, and magenta. Oxygen and nitrogen atoms are shown in red and blue, respectively. Water molecules are shown as spheres in green (panel a), cyan (panel b), and magenta (panel c). Heme irons are shown as orange spheres. (**d**) Overlay of the three strucures.

Supplementary Figure S2 UV-Vis Spectra of reduced and oxidized NwCld. Optical spectra of purified NwCld in the oxidized (solid line) and reduced (dashed line) state. The spectra were recorded in 50 mM potassium phosphate (pH 7.0). The reduced spectrum was measured in presence of 10 mM sodium dithionite.

Supplementary Figure S3 Visualization of the surface of the NwCld holoenzyme. The solvent-accessible surface of the NwCld dimer is colored according to its electrostatic potential (blue for positive, red for negative). Hemes are presented as yellow stick models. The orientation of the structure is the same as in Figure 1a.

Supplementary Figure S4 Visualization of the interface between the two monomers in the NwCld dimer. (**a**) Superposition of an NwCld dimer (colored as in Figure 1a) with an NdCld pentamer (shown in grey), with one monomer of NwCld (cyan) spatially aligned to one NdCld monomer. Note the different location of the second NwCld monomer (yellow) compared to the monomers of NdCld, illustrating differences in the interfaces between the monomers of dimeric and pentameric Clds. (**b**) Detailed view of the NwCld dimer interface. Residues involved in interactions between the two monomers of the NwCld holoenzyme are shown in pale green. Side chains of amino acids involved in the formation of salt bridges are shown as sticks with carbon, oxygen, and nitrogen atoms depicted in grey, red, and blue, respectively. Hemes are presented as orange stick models with heme irons shown as orange spheres. Interactions of heme propionate with the protein backbone are visualized by red dashes.

Supplementary Figure S5 Structure-based amino acid sequence alignment of lineage II Clds (first block of 15 sequences), lineage I Clds (next block of 12 sequences), and Cld-like proteins (remaining six sequences). Names printed in red represent proteins whose structures have been determined and which were used to define the shown alignment. Residues that are conserved in all Clds and Cld-like proteins are colored red. The signature residues, which are conserved only in lineage I and II Clds, are shown in green or cyan (the catalytically important arginine). The published lineage I Cld sequences from *D. chlorophilus*, *P*. sp. PK, *D*. sp. LT-1, *D*. sp. WD, and *P. chloritidismutans* are short due to the PCR primers that were used for *cld* gene amplification (Bender *et al.*, 2004; Cladera *et al.*, 2006). Therefore, the conservation of residue Glu167 and Arg127 (only *P. chloritidismutans*) (NwCld numbering) cannot be verified for these enzymes. Residues marked in yellow are conserved in most lineage II Clds and are involved in the formation of salt bridges at the dimer interface. Note that one of these residues (Asp134) lacks a homolog in lineage I Clds and the other Cld-like proteins, and probably is vital for dimer formation in lineage II Clds. Please refer to Figure 2 for sequence accession numbers.

Ralstonia pickettii (YP_001899051) (Nar/Nap)
plasmid pAKD4 (ADD63334)
Cupriavidus metallidurans (ZP_00599474) (Nar/Nap)
Klebsiella pneumoniae (ABR80622) (Nar/Nap)
Pseudomonas stutzeri (YP_001173821) (Nar/Nap)
Nitrococcus mobilis (ZP_01125763) (Nxr)
Cyanothece sp. PCC 7425 (YP_002482168)[2]
Gloeobacter violaceus (NP_924112)[2]
Pseudomonas aeruginosa (CAH04648) (Nar/Nap)
Limnobacter sp. MED105 (ZP_01913751) (Nap[1])
Marine metagenome read (EBY86119)
Marine metagenome read (9AACY020830705)
Eastern Tropical Pacific read (JCVI_READ_1092301031738)
Bradyrhizobium japonicum (NP_773991) (Nap)
Nitrobacter sp. Nb-311A (ZP_01046929) (Nxr)
Nitrobacter winogradskyi NwCld (YP_319047) (*) (Nxr)
Freshwater sediment metagenome read (ABSP01051938)
Janthinobacterium sp. Marseille (YP_001352597)[2]
Sorangium cellulosum (YP_001615619) (Nap)

Lineage II

Azospira oryzae AoCld (2VXHA) (*)
Dechlorosoma sp. KJ (ACB69921) (*)
Dechloromonas aromatica DaCld (AAZ47311) (*)
Dechloromarinus chlorophilus (AAT07041) (*)
Pseudomonas chloritidismutans (CAI54237) (*)
Pseudomonas sp. PK (AAT07040) (*)
Dechloromonas agitata (AAM92878) (*)
Dechloromonas sp. LT-1 (AAT07039) (*)
Ideonella dechloratans (CAC14884) (*)
Dechlorospirillum sp. WD (AAT07043) (*)
Magnetospirillum magnetotacticum (ZP_00208549)
Candidatus Nitrospira defluvii NdCld (ACE75544) (*)

Lineage I

Blastopirellula marina (ZP_01093420)
Rhodopirellula baltica (CAD76620)
Verrucomicrobium spinosum (ZP_02926541)
Haloferax_volcanii (YP_003535903)
Solibacter usitatus (YP_821642)
Listeria monocytogenes (NP_465637)
Staphylococcus aureus (NP_371111)
Geobacillus stearothermophilus GsCld (1T0TV)
Thermus thermophilus TtCld (1VDHA)
Corynebacterium diphtheriae (NP_939746)
Mycobacterium leprae (CAC31426)
Propionibacterium acnes (YP_055023)
Deinococcus radiodurans (NP_295204)
Rubrobacter xylanophilus (YP_643527)
Herpetosiphon aurantiacus (YP_001546309)
Nitrosopumilus maritimus SCM1 (YP_001581849)
Sulfolobus acidocaldarius (YP_254735)
Thermoplasma acidophilum TaCld (3DTZA)

Supplementary Information: Results and discussion

Active site structure of NwCld

The arginine at position 127 (NwCld numbering) plays an important functional role in Clds as shown for NdCld in a previous study (Kostan *et al.*, in press). Consistently, the side chain conformation of this residue is stabilized through a very similar pattern of interactions in all Cld structures. In NwCld this involves the two water molecules Wat1 and Wat193 (see main text). In the NdCld structure these water molecules are replaced by a molecule of ethylene glycol (originating from the cryo solution) with its oxygen atoms in position of the waters. Instead of interacting with its side chain, the arginine forms two hydrogen bonds to the backbone (Supplementary Figure S1b). The water molecule Wat108 which in NwCld coordinates the iron atom, is in the structure of NdCld replaced by one of the nitrogen atoms of imidazole. The same is true for the structures of AoCld and DaCld, which were determined in complex with thiocyanate and nitrite, respectively: both of these inhibitors replace Wat108. In these two structures the arginine side chain is no longer stabilized by hydrogen bonds to the glutamine side or to the backbone (like in NdCld) but is found in an orientation closer to the iron, with one of its nitrogen atoms replacing Wat95 (Supplementary Figure S1c).

Similar to all Clds and Cld-like proteins whose structures are known, NwCld has a positive electrostatic potential on the surface around the entrance to the active site (Supplementary Figure S3). For chlorite dismutases this is clearly important for attracting the anionic substrate chlorite. In case of the Cld-like proteins this indicates that their substrate, if it is not chlorite, might also be negatively charged.

Dimer interface in NwCld

Until now all structurally characterized Clds and Cld-like proteins are pentamers (Ebihara *et al.*, 2005; Goblirsch *et al.*, 2010; Kostan *et al.*, in press) or, in the case of the Cld from *A. oryzae*, a hexamer (de Geus *et al.*, 2009). The interface between the subunits in all these structures is very similar. Each domain of a subunit interacts with the corresponding half of a neighbouring molecule. The interface consists mainly of residues from the N-terminal helix α4 and strand β4, which interact with residues in the loop between strands β2 and β3 from the N-terminal domain of the neighboring subunit. The second half of the interface consists of the corresponding parts of the C-terminal domain. In contrast, NwCld is a dimer and the interface between monomers is entirely different from that observed in the pentameric or hexameric Clds (Supplementary Figure S4a). Firstly, the NwCld subunit lacks all helices in the N-

terminal domain whereas the loop between strands β2 and β3 is much longer and adopts a completely different conformation. The β2-strand is now part of the second β-sheet interacting with β4', and therefore NwCld lacks half of the interacting surface, which would be needed for pentamerisation. Secondly, the β2 strand in this position along strand β4' would block the interaction between the C-terminal domains. Instead, the dimer interface of NwCld is formed by residues from the loop between β4 and α1' from one subunit which interacts with the loop between helix α2' and helix α3' plus the loop between strand β2' and strand β3' (Supplementary Figure S4b). The surface buried in this interface is 980 Å^2, representing 11% of the total subunit surface. For comparison, the surface buried between two subunits in the pentameric NdCld is 1400 Å^2 corresponding to 11.5 % of the subunit surface (Kostan *et al.*, in press). The interactions involved in NwCld dimer formation are mainly of electrostatic, and to a lesser extent of hydrophobic, nature. Namely, only about 17 % of the area buried in the interface comes from hydrophobic amino acids while the Pisa server (Krissinel and Henrick, 2007) identified 16 inter-domain hydrogen bonds and four salt bridges. One salt bridge is formed between Arg60 located in the loop between β4 and α1' and Asp134 found in the loop between strand β2' and strand β3'. A second bridge connects Arg64 located in the beginning of helix α1' to Glu103 found in the beginning of helix α3'. These two pairs of salt bridges are repeated two times due to the symmetric dimer interface (Supplementary Figure S4b). The residues involved in these salt bridges (Arg60-Asp134 and Arg64-Glu103) are conserved in most Clds of lineage II (Supplementary Figure S5), indicating that these proteins could form a dimer with an interface similar to that of NwCld. In addition, Arg60 interacts with the heme *b via* a hydrogen bond between its backbone amide hydrogen and a propionate group of the heme *b* (Supplementary Figure S4b). Thus, Arg60 is a direct link of interactions from the heme to the dimer interface, which might indicate that formation of the NwCld dimer influences the binding of the heme. This is in contrast to the canonical (lineage I) Clds, which lack any obvious link between the subunit interface and heme binding. The interface of the pentameric NdCld has a similar number of inter-domain hydrogen bonds and salt bridges, but a larger part of the surface has a hydrophobic character (about 30 %).

References

de Geus DC, Thomassen EA, Hagedoorn PL, Pannu NS, van Duijn E, Abrahams JP (2009). Crystal structure of chlorite dismutase, a detoxifying enzyme producing molecular oxygen. *J. Mol. Biol.* **387:** 192-206.

Ebihara A, Okamoto A, Kousumi Y, Yamamoto H, Masui R, Ueyama N *et al* (2005). Structure-based functional identification of a novel heme-binding protein from *Thermus thermophilus* HB8. *J. Struct. Funct. Genomics* **6:** 21-32.

Goblirsch BR, Streit BR, Dubois JL, Wilmot CM (2010). Structural features promoting dioxygen production by *Dechloromonas aromatica* chlorite dismutase. *J. Biol. Inorg. Chem.*

Kostan J, Sjöblom B, Maixner F, Mlynek G, Furtmüller PG, Obinger C *et al* (in press). Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium "*Candidatus* Nitrospira defluvii": Identification of a catalytically important amino acid residue. *J. Struct. Biol.*

Krissinel E, Henrick K (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372:** 774-797.

Supplementary Figure S1
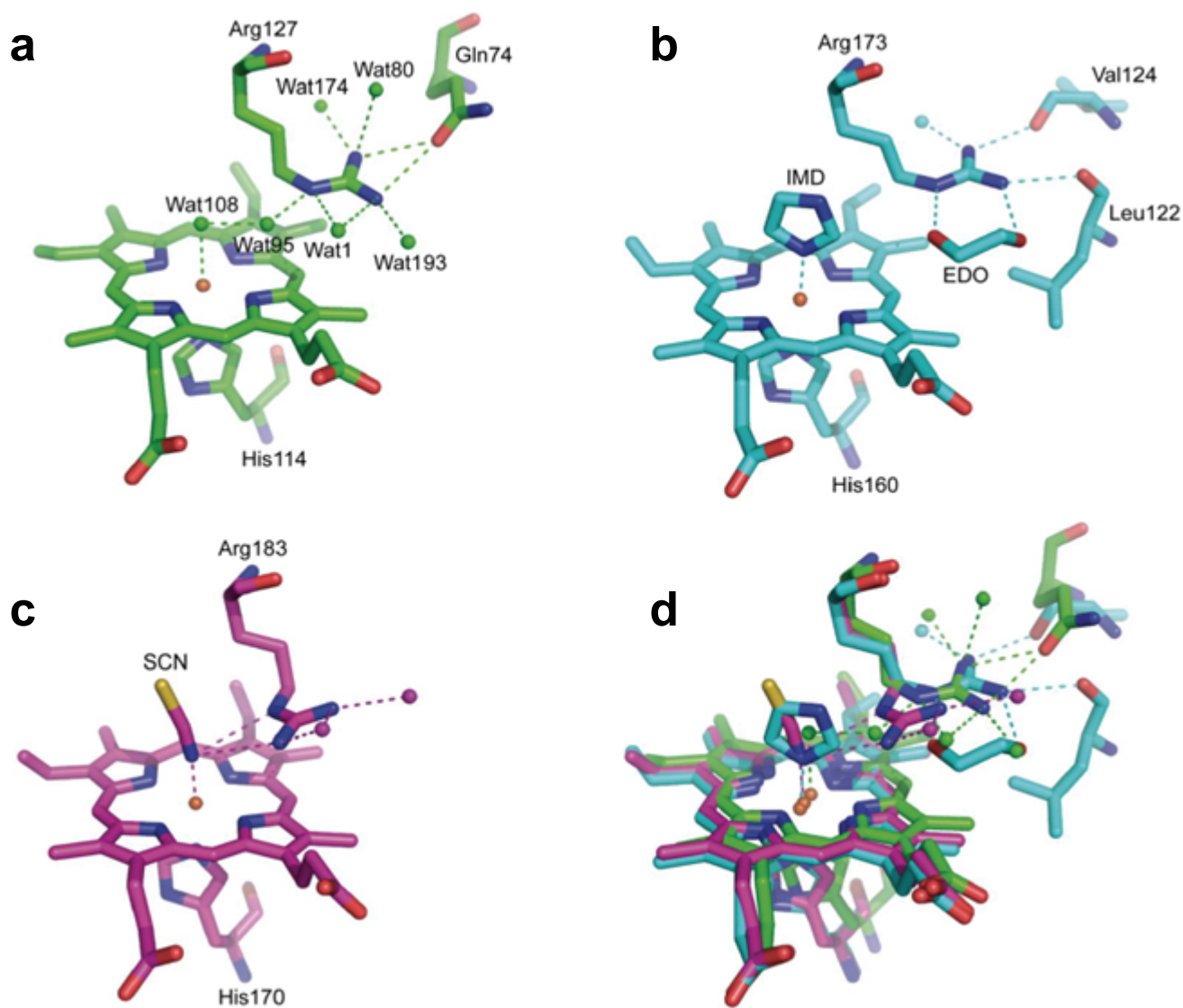
Supplementary Figure S2
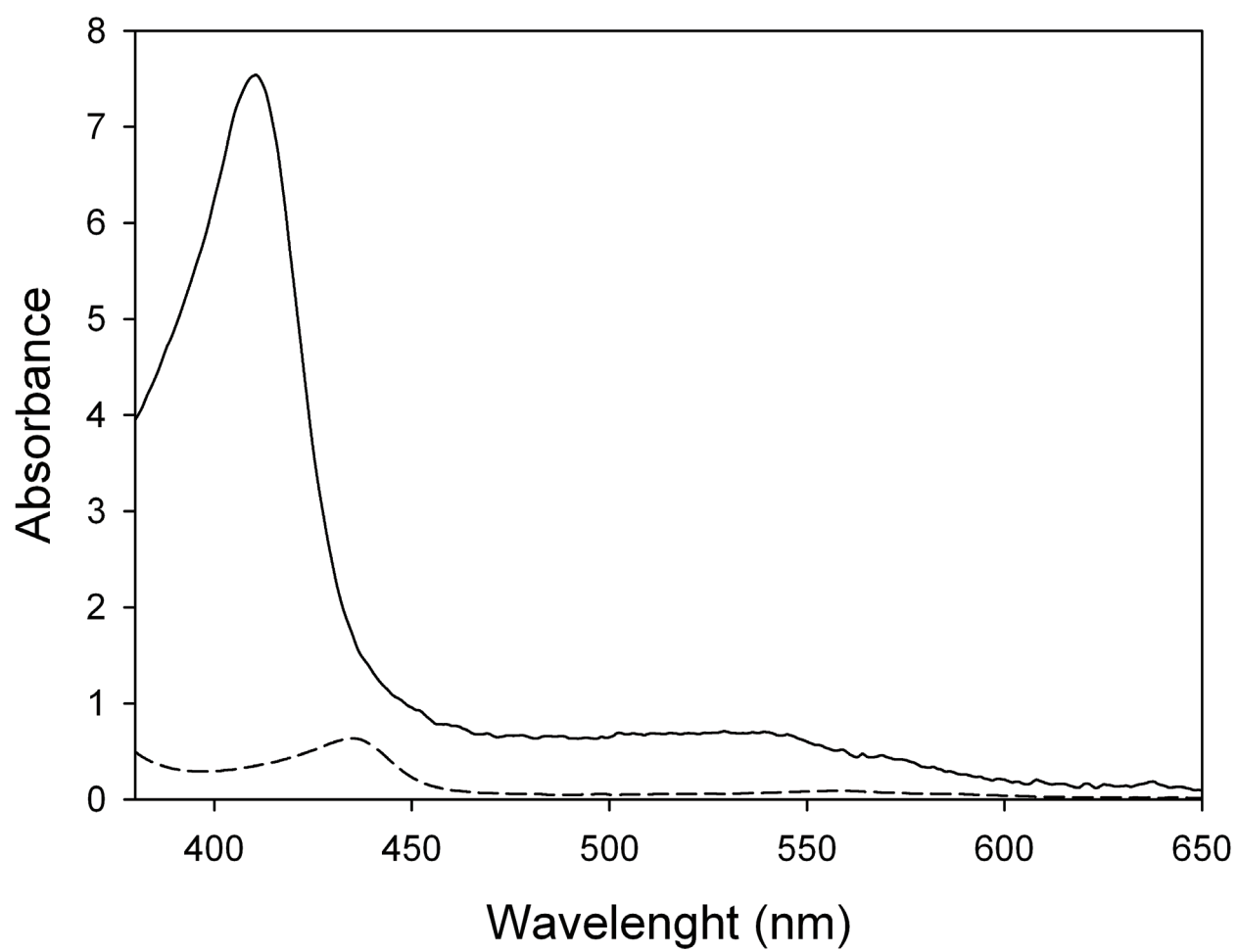
Supplementary Figure S3
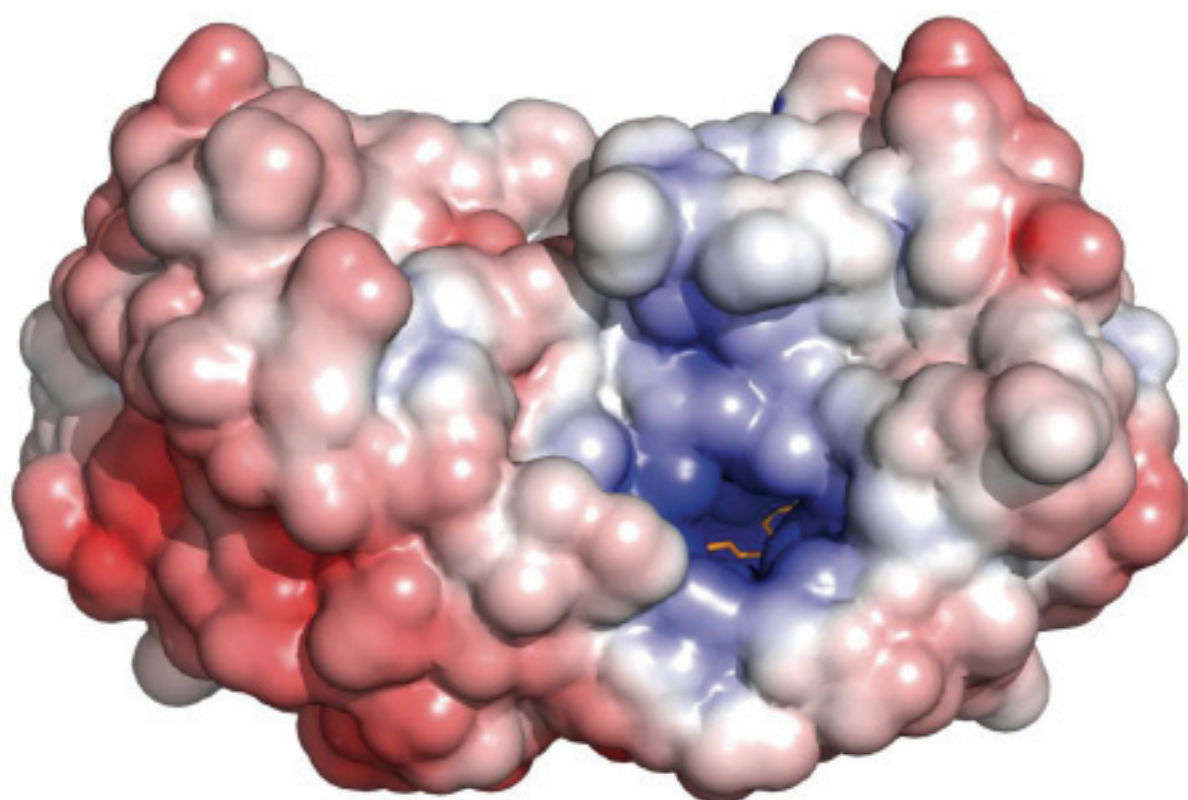
Supplementary Figure S4

a

b
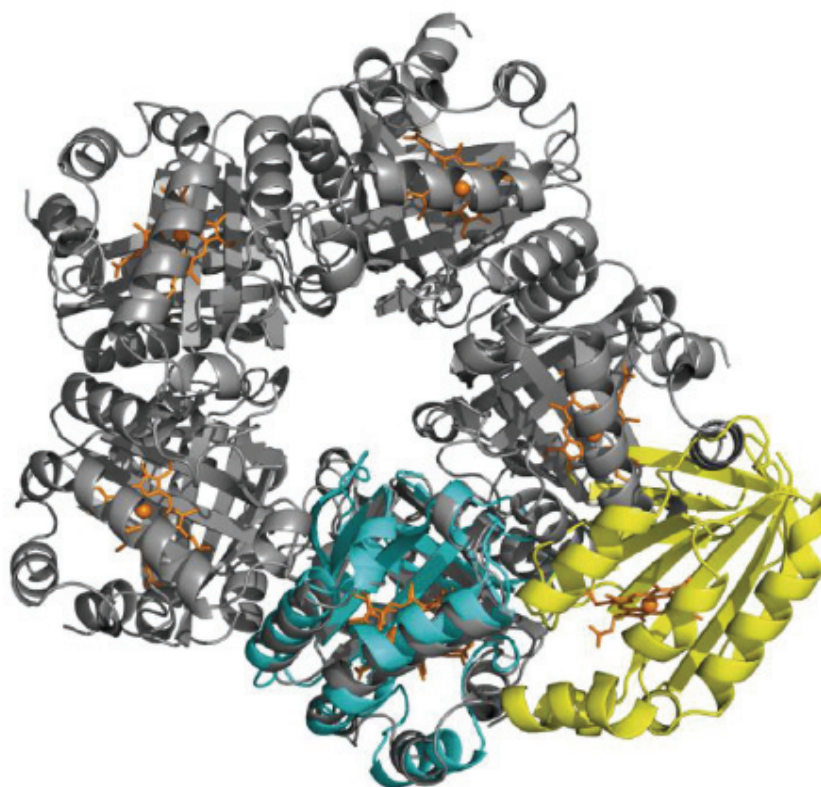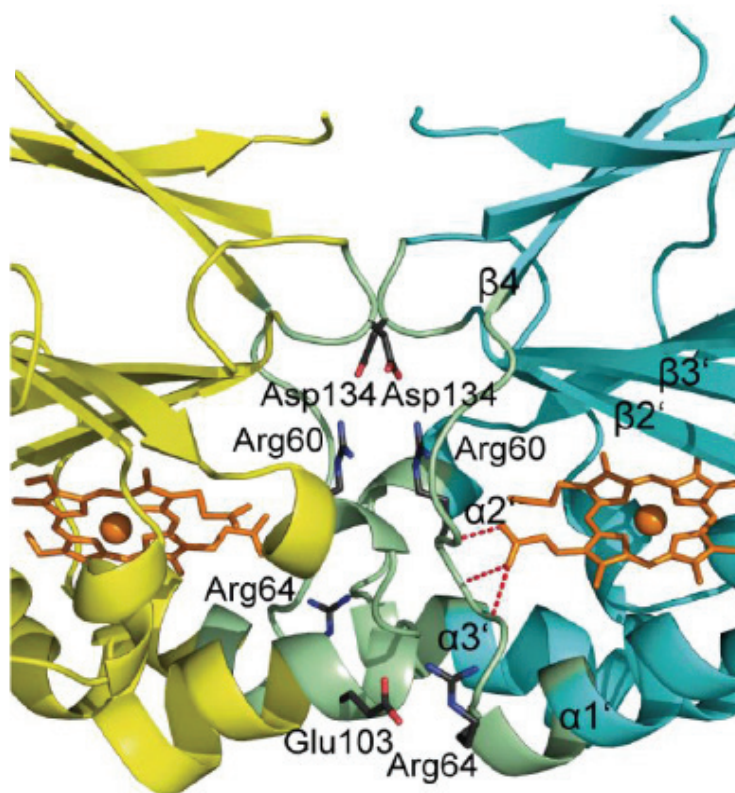
Supplementary Figure S5

```
Nitrobacter winogradskyi              3 -------------FTVFTGGDS-------------------------------------G-AWSI--LSVAPVIGESLMAASHLAIAPSLSLGDTSAT--   47
Nitrobacter sp. Nb-311A               3 -------------FTIFTGGDT-------------------------------------G-AWNI--LSMASVIGDSLMPASHLAIAPSASLGDAAV---   46
Bradyrhizobium japonicum              2 -------------FRTFRGGHS-------------------------------------G-GWRI--TSISPVTGEPLPFMPALSVTDSEAVSLPLVPSR   48
Limnobacter sp. MED105                4 -------------HYSFIGGQE-------------------------------------G-QWRV--TRCDTVVGAPIEAVPRLNVVNT-AASQLSQR--   47
Pseudomonas aeruginosa                4 -------------HYSFIGGSE-------------------------------------G-SWRV--TSCETLIGIPLEIVERVNVVNM-PSTNLIER--   47
Gloeobacter violaceus                 4 -------------RYSFLGGKR-------------------------------------G-PWRV--ARLDGLRGAGLEAVERLQIVQG-EWAESASE--   47
Cyanothece sp. PCC 7425              14 -------------RYSFIGGRT-------------------------------------G-QWQV--VKIRNVLGPGLQLVEKVNILNG-AVAEIPLD--   57
Nitrococcus mobilis                   5 -------------LFAFVGGEI-------------------------------------G-SWRV--IETKTVAGEGLAEVKRLNVVNA-AVPLLPDD--   48
Pseudomonas stutzeri                  3 ----------------------------------------------------------------AVAGAPLPGIPRLNVAAG-SVSPQPPG--         28
Klebsiella pneumoniae                 5 -------------LFTFAGGET-------------------------------------G-VWRV--VRMDAVAGAPLPGIPRLDVAAG-SVSPQPLG--   48
Cupriavidus metallidurans             5 -------------LFAFVGADI-------------------------------------G-PWRI--VRAETRVGEPLEAKRLNVVSA-SELQSETN--   48
Ralstonia pickettii                   5 -------------LFAFVGADI-------------------------------------G-PWRI--VRAETRVGEPLEAKRLNVVSA-SELQSETN--   48
plasmid pAKD4                         0 ..........................................................................................             0
Janthinobacterium sp. Marseille       5 -------------LFGFVGGDS-------------------------------------G-AWEV--TQMRAVKGAPLPEVKTIAILNGFSIQGHHAH--   49
Sorangium cellulosum                  7 -------------RVSFVAGSA-------------------------------------G-AWRV--ERTVLVRGEELPAAPRLQRVEGGSFVAPPEA--   51

Ca. Nitrospira defluvii               1 .ADREKLLTESGVYGTFATFQMDHDWWDLPGESRVIS------VAEVKGLVEQWSGKILVESYLLRGLSD---------HADLMFRVHARTLSDTQQFLS   84
Azospira oryzae                      11 KIERGTILTQPGVFGVFTMFKLRPDWNKVPAMERKGA------AEEVKKLIEKHKDNVLVDLYLTRGLET---------NSDFFFRINAYDLAKAQTFMR   95
Dechloromonas aromatica              42 KIERGTILTQPGVFGVFTMFKLRPDWNKVPVAERKGA------AEEVKKLIEKHKDNVLVDLYLTRGLET---------NSDFFFRINAYDLAKAQTFMR  126
Dechlorosoma sp. KJ                   42 KIERGTILTQPGVFGVFTMFKLRPDWNKVPAMERKGA------AEEVKKLIEKHKDNVLVDLYLTRGLET---------NSDFFFRINAYDLAKAQTFMR  126
Dechloromarinus chlorophilus          1 ...............................................KKLIEKHKDNVLVDLYLTRGLET---------NSDFFFRIHAYDLAKAQTFMR   44
Pseudomonas chloritidismutans         1 ....................................................LVDLYLTRGLET---------NSDFFFRIHAYDLAKAQTFMR          33
Pseudomonas sp. PK                    1 ...............................................KKLIEKHKDNVLVDLYLTRGLET---------NSDFFFRIHAYDLAKAQTFMR   44
Dechloromonas agitata                40 KI-----LTAPGVFGNFSTYKVRPDYYKLSMAERKGA------AAEVVAVVEKYKDKVKAEAYLTRGFEA---------QSDFFLRIHSYDMAATQAFLV  119
Dechloromonas sp. LT-1                5 -------------------------------------------DKHKDKVIVDAYLTRGEA---------NSDYFLRVHSYDMAATQAFLI            44
Ideonella dechloratans               48 KI-----LSAPGVFVAFSTYKIRPDYFKVALAERKGA------ADEVMAVLEKHKEKVIVDAYLTRGYEA---------KSDYFLRVHAYDAVAAQAFLV  127
Dechlorospirillum sp. WD              6 -------------------------------------------KHKDKVIVDTYLTRGLGA---------GSDYLLRVHSTDMAATQAFLV            44
Magnetospirillum magnetotacticum     51 KL-----LTSPGVFGNFSTYKLRSDYYKLSMAERKGA------AAEVVEKHKANIIADAYLTRGFEA---------QSDYFLRVHSMDMAATQAFLV    130

Geobacillus stearothermophilus        6 -------QTLDGWYCLHDFRTIDWSAWKTLPNEEREAAISEFLALVDQWETTESEKQGSHAVYTI--VGQ---------KADILFMILRPTLDELHEIET   87
Listeria monocytogenes                6 -------KTLDGWFCLHDFRSIDWAAWRELNPGNQELMLNELSHFLSDMEITKNIGEGEHTIYSI--LGQ---------KADLVFFTLRDSLEALNEVEN   87
Staphylococcus aureus                 6 -------ETLDGWYSLHLFYAVDWASLRIVPKDERDALVTEFQSFLENTATVRSSKSGDQAIYNI--TGQ---------KADLLLWFLRPEMKSLNHIEN   87

Thermus thermophilus                  3 RHVPEPTHTLEGWHVLHDFRLLDFARWFSAPLEAREDAWEELKGLVREWRELEEAGQGSYGIYQV--VGH---------KADLLFLNLRPGLDPLLEAEA   91

Thermoplasma acidophilum              2 ----------TEIYTSVLSYRLLEGKAYSDADTRSLDR----MMRSIDEFFSANPGYINFHIYRS--YRT---------DSDVIFWYSSRNPDLMILAKE   76
Sulfolobus acidocaldarius             2 ----------ANGVYMYVIQAKFNNEWWSTSLQTRRNI----LNRIEELEARSKNDLVALKRFIS--LRY---------DGHLLYWVSDFDTSKLNNLRY   76


Nitrobacter winogradskyi             48 ------------TPWQLRGVASHARYVERAE---KIALT----SVQAGLGRNEATRAALIPIRKSA-----AKWEMTQDERRAIFEDKSHHIAASLKYLP-  123
Nitrobacter sp. Nb-311A              47 ------------TPWRLRGVTSHLRYVERAE---KIALT----EVQAGLGRAEATRAALIPIRKSA-----AKWELTQDERRKVFEDKSHHIAASLKYLP-  122
Bradyrhizobium japonicum             49 ------------NAWRLAGVPSSLRYTERAE---KQQLV----AVQAGLGRLEATSAALIPIRKSQ-----AKWELTQEERRIFEDRSHHIASSLRFLP-  124
Limnobacter sp. MED105               48 ------------GTWMLQGFTSNVRYAERHE---INQLR----AKQEGLSRPASTCAALIPIKKNA-----QQWALSQDERRAIFEAQSHHTEIGLAYLP-  123
Pseudomonas aeruginosa               48 ------------GTWVLQGFTSNVRYAERHE---INQLR----AKQEELNRPTSSCAALIPIKKSP-----AKWAMSQDERREIFEAKSHHTEIGLAYLP-  123
Gloeobacter violaceus                48 ------------AAWVLRGLTSNVRYATRPE---VDALR----ERQPALARPEARCAALIPIKKSA-----RKWELAQDERRAIFEETSHHTAIGMEFLP-  123
Cyanothece sp. PCC 7425              58 ------------SAWRLQGFASNIRYAIRTE---LEALQ----AVQPMLNRAEAILAVLIPIKKSA-----QQWEMAQDERRDIFERESHHTAVGLEYLP-  133
Nitrococcus mobilis                  49 ------------AQWLLRGVTSNERYVTRSE---RAQLT----AKQPVLGRRQATCAAFIPIRKTA-----SKWNLAQDERRMILEESSNHIKTGLKYLP-  124
Pseudomonas stutzeri                 29 ------------TKWLLRGITSNERYVREE---KDRLV----AKQPSLGRAEATCAALIPIRKNP-----SKWGLAQDERRKIFEEQSRHIHIGLQYLP-  104
Klebsiella pneumoniae                49 ------------TKWLLRGITSNERYVVREE---KDRLV----AKQPSLGRAEATCAALIPIRKNP-----SKWGLAQDERRKIFEEQSRHIHIGLQYLP-  124
Cupriavidus metallidurans            49 ------------APWILRGITSNERYVMRAE---KNEIV----AKQQGLARPEATCGALIPIRKNA-----AKWELTQDERRSVFE-QSKHVQIGLQYLP-  123
Ralstonia pickettii                  49 ------------APWILRGITSNERYVMRAE---KNEIV----AKQQGLARPEATCGALIPIRKNA-----AKWELTQDERRSVFE-QSKHVQIGLQYLP-  123
plasmid pAKD4                         1 .........................MRAE---KNEIV----AKQQGLARPEATCGALIPIRKNA-----AKWELTQDERRSVFE-QSKHVQIGLQYLP-   60
Janthinobacterium sp. Marseille      50 ------------WVLRGVTSNERYVTKEE---KSRLI----ATQEGLGRTESTLAALIPIRKNA-----SKWLLTQEERREILEERSHHIQIGMAYLP-  123
Sorangium cellulosum                 52 ------------TWVLGGVRSNERYVELEE---RRRLV----AVQEDLGRASSTQAALLPLRKSK-----AKWALAQDERRAIFEARSRHIEIGLEYLP-  126

Ca. Nitrospira defluvii              85 AFMGTRLGRHLTSGGLLHGVSKKPTYVAGFP---ESMKT----ELQVNGESGSRPYAIVIPIKKDA-----EKWALDQEARTALMQ---EHTQAALPYLK-  169
Azospira oryzae                      96 EFRSTTIGKNADVFETLVGVTKPLNYISKDK---SPGLNA--GLSSATYSGPAPRYVIVIPVKKNA-----EKWNMSPEERLKEME---VHTTPTLAYLV-  182
Dechloromonas aromatica             127 EFRSTTVGKNADVFETLVGVTKPLNYISKDK---SPGLNA--GLSSATYSGPAPRYVIVIPVKKNA-----EKWNMSPEERLKEME---VHTTPTLAYLV-  213
Dechlorosoma sp. KJ                 127 EFRSTTIGKNADVFETLVGVTKPLNYISKDK---SPGLNA--GLSSATYSGPAPRYVIVIPVKKNA-----EKWNMSPEERLKEME---VHTTPTLAYLV-  213
Dechloromarinus chlorophilus         45 EFRSTTVGKNADVFETLVGVTKPLNYISKDK---SPELNA--GLSSATYSGPAPRYVIVIPVKKNA-----EKWNMSPEERLKEME---VHTTPTLAYLA-  131
Pseudomonas chloritidismutans        34 EFRSTTVGKNADVFETLVGVTKPLNYISKDK---SPELNA--GLTSATYSGPAPRYVIVIPVKKNA-----EKWNMSPEERLKEME---VHTTPTLAYLV-  120
Pseudomonas sp. PK                   45 EFRSTTVGKNADVFETLVGVTKPLNYISKDK---SPELNA--GLSSATYSGPAPRYVIVIPVKKNA-----EKWNMSPEERLKEME---VHTTPTLAYLV-  131
Dechloromonas agitata               120 DFRATRFGMNAEVTENLVGMTKDLNYITKDK---SPNLNA--GLTGATYRDATPRYAFVIPVKKSA-----DKWNLTDEQRLKEME---THTLPTLANLV-  206
Dechloromonas sp. LT-1               45 DFRATRFGMYSDVTEQLLGMTKPLNYITKEK---SPDLNA--GLSSATYSAEAPPYAFVIPVKKSA-----EKWNLTDAQRLKEME---THTLPTLGNLV-  131
Ideonella dechloratans              128 DFRATRFGMYSDVTELVGITKALNYISKDK---SPDLNK--GLSGATYAGDAPRFAFMIPVKKNA-----DKWNLTDEQRLKEME---THTLPTLPFLV-  214
Dechlorospirillum sp. WD             45 DWRATKLGMYSDVTENLVGITKALNYISKDK---SPELNA--GLSSATYSDSAPRYVIVIPVKKNA-----DKWNMSDEQRLKEIE---VHTHGTLQYLV-  131
Magnetospirillum magnetotacticum    131 DFRATRFGMYSDVTENLVGITKALNYISKDK---SPDLNS--GLSSASYTGDAPRYAIMIPVKKDA-----AKWNRSDEERLKEME---VHTQPTLQYLV-  217

Geobacillus stearothermophilus      88 ALNKTKLADYLLPAYSYVSVVELSNYLASGS---EDPYQIPEVRRRLYPILPKTNYICFYPMDKRRQGNDNHYMLSMEQRRELMR---AHGMTGRKYAG-  180
Listeria monocytogenes              88 RFNKLAIADYLLPTYSYISVVELSNYLASHMAGGDDPYQNKGVRARLYPALPPKKHICFYPMSKKRDGADNHYMLPMEERQQLIR---DHGLIGRSYAG-  183
Staphylococcus aureus               88 EFNKLRIADFLIPTYSYVSVIELSNYLAGKSD---EDPYENPHIKARLYPELPHSDYICFYPMNKRRNETYNHYMLTMEERQKLMY---DHGMIGRKYAG-  181

Thermus thermophilus                92 RLSRSAFARYLGRSYSFYSVVELGSQEKPLD---PESPY---VKPRLTPRVPKSGYVCFYPMNKRRQGQDNHYMLPAKERASLMK---AHGETGRKYQG-  181

Thermoplasma acidophilum            77 R-VQASMRPIAVSSFSSISIYDESPYNAMNK---KL---------EDSLRLPPLRYFVAYPMSKTP----DYYLLDFDTRKEIMH---EHIKMALNHPDE  156
Sulfolobus acidocaldarius           77 S-LISSGEGFLEEKLTLFSYFKPSPYIGGSAD---KL---------ASYLRLEPLRYFIAYPMKKSP----RKYLLPFEERKEIMD---EHIEIAKTHPDN  157


Nitrobacter winogradskyi           124 -AIARQLYHCRDIG---EPFDFLTWFEYAPEHATMFEDLVGVLRATEEWTY--VEREVDIRLAR-AI..........   183
Nitrobacter sp. Nb-311A            123 -AIARQLYHCRDLG---EPFDFLTWFEYAPEHATMFEDLVGVLRATEEWTY--VEREVDIRLAR-AV..........   182
Bradyrhizobium japonicum           125 -AIARQLYHCRDLG---GPFDFLTWFEFAPAHASLFEELVAMLRRTEEWTY--VEREVDVRVVK-EVLSA........   187
Limnobacter sp. MED105             124 -EIARQLHHSRDLG---EPFDFLTWFEFAPEHTDAFDELLVKLRTSEEWEY--VEREVDIRLVK-DSL........   184
Pseudomonas aeruginosa             124 -EIARQLHHSRDLG---EPFDFLTWFEFAPEHTAIFNKLLAQLRSSKEWEY--VEREIDIRLVK-NV........   183
Gloeobacter violaceus              124 -AVARRLHHCRDIG---EPFDFLTWFEYAPEHTHAFEDLLDRLRVTREWDF--VEREVDIRLEH-IDARI.......   186
Cyanothece sp. PCC 7425            134 -GVARRLLHCRDLG---EEFDFLTWFEFAPEHSSAFNELLLRMRASKEWEY--VEREVEVWLKR-L........   192
Nitrococcus mobilis                125 -AVARRLHHCRDLGGDAEPFDFLTWFEYAPSDSAAFDELVAELRASQEWVY--VDREIDMRLAR-DE........   187
Pseudomonas stutzeri               105 -AVARRLHHCRDLG---ENEPFDFLTWFEYSPSDETAFNRLLAELRASEWQY--VDREIDIRLVH-EPA........   168
Klebsiella pneumoniae              125 -AVARRLHHCRDLG---ESEPFDFLTWFEYSPSDEPGFNRLLAELRASVEWKY--VDREIDIRLVH-EPA........   187
Cupriavidus metallidurans          124 -AVARKLHHCRDLS---ENEPFDFLNWFEYAPIHEVENRLLSELRASEEWKY--VDREVDIRLTQ-AQV........   186
Ralstonia pickettii                124 -AVARKLHHCRDLS---ENEPFDFLNWFEYAPIHEVENRLLSELRASEEWKY--VDREVDIRLTQ-AQV........   186
plasmid pAKD4                       61 -AVARKLHHCRDLS---ENEPFDFLNWFEYAPIHEVENRLLSELRASEEWKY--VDREVDIRLTQ-AQV........   123
Janthinobacterium sp. Marseille    124 -AIARRLYHCRDIG---TPEPFDFLTWFEFSPSDASQFDDLVAQLRTTHEWSY--IDHEIDIRLAR-KNG........   186
Sorangium cellulosum               127 -AVARRLYHSRDLG---GPFDFLTWFEFAEGDAGAFDELVGRLRETEEWSY--VEREIEVRLSR-A...........   185

Ca. Nitrospira defluvii            170 -TVKRKLYHST-GL---DDVDFITYFETE--RLEDFHNLVRALQQVKEFRHNRRFGHPTLLGTM-SPLDEILEKFAQ   238
Azospira oryzae                    183 -NVKRKLYHST-GL---DDTDFITYFETD--DLTAFNNLMLSLAQVKENKFHVRWGSPTTLGTI-HSPEDVIKALAD   251
Dechloromonas aromatica            214 -NVKRKLYHST-GL---DDTDFITYFETD--DLTAFNNLMLSLAQVKENKFHVRWGSPTTLGTI-HSPEDVIKALAD   282
Dechlorosoma sp. KJ                214 -NVKRKLYHST-GL---DDTDFITYFETD--DLTAFNNLMLSLAQVKENKFHVRWGSPTTLGTI-HSPEDVIKALAD   282
Dechloromarinus chlorophilus       132 -NVKRKLYHST-GL---DDT.......................................................   146
Pseudomonas chloritidismutans      121 -NV........................................................................   122
Pseudomonas sp. PK                 132 -NVKRKLYHST-GL---DDT.......................................................   146
Dechloromonas agitata              207 -NVKRKLYHST-GL---DDTDFITYFETA--DLGAFNNLMLALAKVPENKYHVRWGSPTVLGTI-QSFDSVVNTLSM   275
Dechloromonas sp. LT-1             132 -NVKRKLYHST-GL---DDT.......................................................   146
Ideonella dechloratans             215 -NVKRKLYHST-GL---DDTDFITYFETN--DLGAFNNLMLSLAKVPENKYHVRWGNPTVLGTI-QPIENLVKTLSM   283
Dechlorospirillum sp. WD           132 -NVKRKLYHST-GL---ADA.......................................................   146
Magnetospirillum magnetotacticum   218 -NVKRKLYHST-GL---DDTDFITYFETN--DLGAFNNLLISLAKVPENKHHVRWGSPTVLGTI-QTFETVVKTLSM   286

Geobacillus stearothermophilus     181 -KVTQIITGSV-GL---DDFEWGVTLFSD--DALQFKKLVYEMRFDEVSAR-FGEFGSFFVGTRLPMENVSSFFHV.   248
Listeria monocytogenes             184 -KVQQIIGGSI-GF---DDYEWGVTLFSD--DALEFKRIVYEMRFDEASAR-YAEFGSFFIGNLLLSEQLSKLFTI.   251
Staphylococcus aureus              182 -KIKQFITGSV-GF---DDFEWGVTLFSD--DVLQFKKIVYEMRFDETTAR-YGEFGSFFVGHLINTNEFDQFFAIS   250

Thermus thermophilus               182 -EVMQVISGAQ-GL---DDWEWGVDLFSE--DPVQFKKIVYEMRFDEVSAR-YGEFGPFFVGKY-LDEEALRAFLGL   249

Thermoplasma acidophilum           157 KGIRSYTTYSF-GI---GDQEFVVLYEIP--DIAAWSRVTEKLREARARKW--IIKETPILLGRL-VDAGDIAGFLL.   224
Sulfolobus acidocaldarius          158 QGIRSYTTYSF-GI---ADYEFVVIYEAP--DLSKWINVVERLREAKARKW--VVSEEPILVGE----IGSLDIFLK.   222
```

Table S1. Data collection and refinement statistics for the NwCld crystal structure.

A. Data Collection

Beamline	ID14-4 (ESRF)
Wavelength (Å)	1.278
Resolution (Å)[1]	45.8 - 2.1
	(2.15 - 2.10)[1]
Space group	$P4_1$
Unit cell (Å, º)	a = b = 102.62,
	c = 49.12
Molecules / a.u.	2
Unique reflections	54986 (3612)
Total reflections obs	2080683 (132332)
Completeness (%)	94.0 (83.1)
R_{meas}[2]	0.081 (0.769)
R_{mearge}[3]	
Multiplicity	37.8 (36.6)
I/σ(I)	39.38 (7.43)

B. Phasing

No. of sites	2
Figure of merit acentric (centric)	0.308 (0.176)

C.Refinement

R_{cryst}[4]/ R_{free}[5]	0.191/0.239
R.m.s.d. bonds (Å)	0.012
R.m.s.d. angles (º)	1.314

[1] Values in parentheses are for the highest resolution shell.

$$^2 \; R_{meas} = \frac{\sum_h \sqrt{\frac{n_h}{n_h - 1} \sum_i^{n_h} \left| \hat{I}_h - I_{h,i} \right|}}{\sum_h \sum_i^{n_h} I_{h,i}} \text{ with } \hat{I}_h = \frac{1}{n_h} \sum_i^{n_h} I_{h,i}$$

$$^3 \; R_{merge} = \frac{\sum_h \sum \left| \hat{I}_h - I_{h,i} \right|}{\sum_h \sum_i I_{h,i}}$$

Where I (hkl) is the mean intensity of multiple I_i (*hkl*) observations of the symmetry-related reflections, N is the redundancy, n_h is the multiplicity, $\hat{I}_h$ is the average intensity and $I_{h,i}$ is the observed intensity.
[4] $R_{cryst} = \Sigma |F_o - F_c| / \Sigma F_o$
[5] R_{free} is the cross-validation R_{factor} computed for the test set of reflections (5%) which are omitted in the refinement process.

Table S2. Coverage of lineage I chlorite dismutase genes by published PCR primers.

Primers targeting *cld*-genes

● = Full match of the primer to the target gene sequence.

○ = Primer has base mismatches to the target gene sequence, or no target sequence information is available at the primer binding site, but successful amplification of the target gene by using this primer was reported in the literature.

Empty fields indicate that the primer has mismatches to the target sequence and amplification of the target gene by using this primer was not reported.

	DCD-F[1]	DCD-R[1]	UCD-238F[1]	UCD-646R[1]	ICD-741F[2]	ICD-1140R[2]	CD-79F[2]	CD-831R[2]	CD-PX1[2]	CD-441F[2]	CD-834R[2]	F[3]	R[3]	CLDF[4]	CLDR[4]
Azospira oryzae	○	○	●	○											
Dechlorosoma sp. KJ			●	●											
Dechloromonas aromatica	●	○	●	●											
Dechloromarinus chlorophilus	○	○	●	●											
Pseudomonas chloritidismutans	○	○	○	○											
Pseudomonas sp. PK	○	○	●	●											
Dechloromonas agitata	●	●	○	○			●	●	●	●	○				
Dechloromonas sp. LT-1	○	○	●	●											
Ideonella dechloratans	●	●	○	●	●	○						○	○		
Dechlorospirillum sp. WD	○	○	●	●											
Magnetospirillum magnetotacticum	●	●	○	●											
Ca. Nitrospira defluvii		●												●	●

[1] Bender *et al.* (2004)

[2] Bender *et al.* (2002)
[3] Thorell *et al.* (2002)
[4] Maixner *et al.* (2008)

References

Bender KS, O'Connor SM, Chakraborty R, Coates JD, Achenbach LA (2002). Sequencing and transcriptional analysis of the chlorite dismutase gene of *Dechloromonas agitata* and its use as a metabolic probe. *Appl. Environ. Microbiol.* **68:** 4820-4826.

Bender KS, Rice MR, Fugate WH, Coates JD, Achenbach LA (2004). Metabolic primers for detection of (Per)chlorate-reducing bacteria in the environment and phylogenetic analysis of *cld* gene sequences. *Appl. Environ. Microbiol.* **70:** 5651-5658.

Maixner F, Wagner M, Lücker S, Pelletier E, Schmitz-Esser S, Hace K *et al* (2008). Environmental genomics reveals a functional chlorite dismutase in the nitrite-oxidizing bacterium '*Candidatus* Nitrospira defluvii'. *Environ. Microbiol.* **10:** 3043-3056.

Thorell HD, Karlsson J, Portelius E, Nilsson T (2002). Cloning, characterisation, and expression of a novel gene encoding chlorite dismutase from *Ideonella dechloratans. Biochim. Biophys. Acta* **1577:** 445-451.

Chapter IV

Biochemical, structural and functional analyses of the chlorite dismutase enzyme from *Listeria monocytogenes*

Manuscript in preparation

Introduction

Listeriae are gram-positive, non-sporulating, rod shaped, motile, facultative-anaerobic bacteria occurring ubiquitously in soil and water that can grow at a broad range of pH (pH 4.5-9.0) and at temperatures between 0°C and 45°C (Hain *et al.*, 2007). The genus Listeria belongs to the phylum Firmicutes which also comprises the genera Bacillus, Clostridium, Enterococcus, Streptococcus, and Staphylococcus.

The six species of the genus Listeria include *Listeria monocytogenes, Listeria ivanovii, Listeria innocua, Listeria seeligeri, Listeria welshimeri*, and *Listeria grayi (Schmid et al., 2005)*, but only two of these - *L. monocytogenes* and *L. ivanovii* are pathogenic. *L. monocytogenes* is an intracellular, foodborne pathogen potentially lethal for humans and animals, whereas *L. ivanovii* is pathogenic just in animals.

The virulence mechanisms have been intensively studied and it is confirmed that the pathogenicity of these two Listeria species is enabled by an approximately 9 kb virulence gene cluster not found in the other Listeria species (Schmid *et al.*, 2005).

The disease caused by *Listeriae* is called listeriosis and 99% of all human listeriosis cases are caused by consumption of contaminated food products (Mead *et al.*, 1999). After ingestion *L. monocytogenes* is able to cross the intestinal, the blood-brain and the placental barrier. Human listeriosis usually manifests as meningitis, meningoencephalitis, septicemia, abortion, prenatal infection and also gastroenteritis (Drevets and Bronze, 2008).

Although listeriosis with 2–15 cases per million population per year seems not to be a great problem, the high mortality rate of about 20–30% in individuals at risk developing listeriosis (pregnant women, elderly, and immunocompromised persons) makes *L. monocytogenes* a serious human pathogen (Mead *et al.*, 1999). A recent outbreak through contaminated cheese recently caused eight deaths in Austria (http://derstandard.at/1267132057525/Verseuchter-Quargel-Listerien-Acht-Tote-wenig-Information).

The finding that *L. monocytogenes* and other intracellular bacteria carry a gene with high sequence similarity to chlorite dismutases (Clds) (Maixner *et al.*, 2008) found in perchlorate reducing bacteria (PCB) was quite surprising, especially because the presence of Clds was thought to be restricted only to organisms capable of energy gaining respiration of (per)chlorate (Coates and Achenbach, 2004). Reduction of perchlorate in PCB leads to formation of chlorate that is further transformed to chlorite, in reaction catalyzed by perchlorate reductase (Streit and DuBois, 2008).

Interestingly, a gene encoding perchlorate reductase was not found in *Listeriae*, indicating that Cld could have different function (than chlorite reduction) in these pathogens. It was speculated that Cld in Listeriae might be involved in scavenging of reactive oxygen intermediates (ROI) and/or reactive

LmCld is expressed constitutively during growth at 37°C as a protein of about 29 kDa (Füreder, 2009). Further analysis showed that the protein is localized in the cytosol which is in accordance with the lack of a signal peptide in the N-terminal part of Lmo2113 (REF). In comparison, validated chlorite dismutases of the PCB *Ideonella dechloratans* (Stenklo *et al.*, 2001), *Azospira oryzae* (de Geus *et al.*, 2009) and *Dechloromonas aromatica* (Goblirsch *et al.*, 2009) were found in the periplasm. The validated Cld of the nitrifier "*C. N. defluvii*" defluvii (NdCld) is also found in the periplasm whereas the Cld of *Nitrobacter winogradsky* (NwCld) is found in the cytoplasm.

No functional or structural analysis of the chlorite dismutase from *Listeria monocytogenes* (LmCld) has been performed to date. We set for a series of biochemical and structural studies with the aim to reveal the functional role of this enzyme in *Listeria monocytogenes*.

Materials and Methods

Heterologous expression and purification of LmCld

Recombinant *E. coli* C43(DE3) strain (Miroux and Walker, 1996) carrying plasmid pET21b(+) (Novagen) containing the coding sequence of the chlorite dismutase gene (ORF annotated as *lmo2113* in EGD-e genome, accession no. NC003210*)* was kindly provided by Stephanie Füreder (Füreder, 2009).

For heterologous expression, Luria-Bertani (LB) medium supplemented with carbenicillin (100 μg/ml) and hemin (50 μg/ml, Fluka) was inoculated with a freshly prepared overnight culture (at a dilution ratio of 1:100). The culture was grown at 37°C under agitation (220 rpm) to mid-log phase (OD_{600}=0.8). For expression induction, temperature as well as the shaker speed were reduced to 25°C and 180 rpm, respectively, 30 minutes before adding isopropyl-β-D-thiogalactopyranoside (IPTG) to the final concentration of 0.1 mM. After 12 hrs, the culture was centrifuged and the resulting cell pellet was frozen in liquid nitrogen and stored at -80°C for further use.

The thawed cell pellet from 1 liter of expression media was resuspended in 35 ml lysis buffer (50 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.4, 250 mM NaCl, 5% glycerol, 1 mM PMSF (phenylmethylsulfonylfluorid), 142 μM hemin, 20 mM imidazole,) and subjected to sonication. Subsequently, cell debris was removed by centrifugation (45000 g, 20 min, 4°C).

The sample was loaded on a 20 ml HisTrap FF crude column (GE Healthcare) equilibrated with solution A (50 mM HEPES pH 7.4, 150 mM NaCl, 20 mM imidazole, 2% glycerol). After washing of unbound protein with the solution A containing 40 mM imidazole, bound protein was eluted with step gradient of 500 mM imidazole in solution A.

nitrogen species (RNI) and thus contributes to a certain degree to resistance against oxidative stress in infected phagocytic cells (Füreder, 2009).

Genomic analysis of the LmCld gene (*lmo2113*) in *Listeria monocytogenes* was performed by Stephanie Füreder (Füreder, 2009). Part of the text from her Ph.D. thesis(Füreder, 2009) was slightly annotated and inserted here to complement our study:

> "In the available sequenced genome of *L. monocytogenes* strain EGDe, the 756 bp-long coding sequence of the cld-like gene (*lmo2113*) is located downstream of two genes involved in sugar metabolism (lmo2108, lmo2110), a gene encoding a putative hydrolase/acyltransferase belonging to the esterase/lipase superfamily (lmo2109) and a gene encoding a putative FMN (flavin mononucleotide)-containing NADPH-linked flavin/nitro reductase gene (*lmo2111*).
>
> Downstream of the cld homologue are genes encoding a putative ATP-binding protein (*lmo2114*) and a putative permease (*lmo2115*), both belonging to the ABC transporter family, as well as two genes coding for hypothetical proteins of unknown function (*lmo2116*, *lmo2117*). Interestingly, the gene *lmo2112* located directly upstream of *lmo2113* shows homology to a putative DNA-binding domain of the excisionase family. Excisionase-like genes usually are located on temperate phages, plasmids or transposons, and are involved in excisive genomic recombination (Swalla *et al.*, 2003). The described lmo2113 neighbourhood is conserved in the five complete genome sequences of *L. monocytogenes* and *L. innocua* available to date. The only published complete genome sequence of *L. welshimeri* (strain SLCC 5334), however, lacks the genes lmo2108 and lmo2109.In the draft genome of *L. grayi*, only a homologue to the cld-like gene *lmo2113* was detected, whereas homologues to any other genes in the region between *lmo2108-lmo2117* were not found in the available sequence data from this organism. Based on the combined occurrence of homologues to *lmo2112* and *lmo2113* in all sequenced Listeria spp., it is tempting to speculate that mobile genetic elements might have been involved in the acquisition of the cld homologue by this genus. Comparative analysis of the region around *lmo2113* to the genomic context of cld-like genes in other organisms than Listeria did not reveal any shared synteny and thus provided no clue regarding similar functions of LmCld and other Cld-like proteins. Consistent with this observation, the Cld-like proteins from different bacteria and archaea display only low sequence similarity, indicating a functional diversification within this protein family (Maixner *et al.*, 2008)".
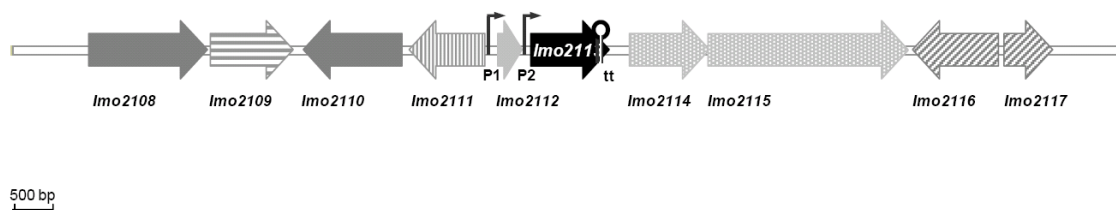

500 bp

Figure 1. with permission of Stephanie Füreder. **Genetic organization of the *lmo2113* locus in *L. monocytogenes*.** Schematic representation of the genomic neighbourhood of the identified *cld*

homologue, *lmo2113*, in *L. monocytogenes.* Putative functions of the adjacent genes are as follows: *lmo2108*, putative N-acetylglucosamine-6-phosphate deacetylase; *lmo2109*, putative hydrolase or acyltransferase; *lmo2110*, putative mannose 6-phosphate isomerase; *lmo2111*, putative FMN-containing NADPH-linked nitro/flavin reductase; *lmo2112*, putative DNA binding domain of excisionase family; *lmo2113*, *cld* orthologue; *lmo2114*, putative ABC transporter (ATP-binding protein); *lmo2115*, putative ABC transporter (permease); *lmo2116*, hypothetical protein; *lmo2117*, putative acyltransferase. Arrows indicate the direction of translation of the genes. P1 and P2, predicted promoters of *lmo2113*. tt, predicted transcription terminator of *lmo2113*. The gene annotation and names refer to those of the genome sequence of *L. monocytogenes* strain EGD-e (accession no. NC003210). The genome region is drawn to scale.

Phylogeny and sequence analyses

Füreder et al. found that *Bacillales* Cld-like proteins cluster together and are clearly distinct from the validated chlorite dismutases of PCB and from "*C. N. defluvii*". The Cld-like protein of the Deinococcus-Thermus phylum member *Thermus thermophilus* (Ebihara *et al.*, 2005) is relatively closely related to the Cld-like sequences of *Bacillales*, indicating lateral gene transfer and possibly also similar functions of these proteins (Füreder, 2009).

The genome sequences of *L. monocytogenes* and the non-pathogenic species *L. innocua* were published by the European Listeria Genome Consortium in 2001 (Glaser *et al.*, 2001). It was shown that *lmo2113* is identical in *L. monocytogenes* and *L. innocua* suggesting that it is not a virulence gene. Additionally Glaser et al. compared the two Listeria genomes with the one from Bacillus. subtilis and suggested a common origin for *B. subtilis*, *L. monocytogenes* and *L. innocua*. As *B. subtilis* is an intensively studied microorganism, comparisons of the *cld* genes and expression profiles were explored. In *B. subtilis,* a significant upregulation of the Cld-homolog (encoded by *ywfI gene)* under different anaerobic growth conditions was shown by Marino et al. (Marino *et al.*, 2000). Similarly, expression levels of LmCld were found to be almost 5 times higher under anaerobic conditions compared to the aerobic conditions (Füreder, 2009).

In contrast, neither Camejo et al. (Camejo *et al.*, 2009) using a DNA macroarray technology to profile the transcriptome of Listeria during mouse infection nor (Toledo-Arana *et al.*, 2009) using tiling arrays and RNAs from wild-type and mutant bacteria grown in vitro, *ex vivo* could detect an up- or downregulation of *lmo2113* in *L. monocytogenes*.

In the Ph.D. thesis of Stephanie Füreder strategies to produce a *lmo2113*-deficient mutant of *L. monocytogenes* strain L028 failed, leading to the conclusion that *lmo2113* is an essential gene (Füreder, 2009). This correlates with a study of Ji et al. (Ji *et al.*, 2001) who applied a genome-wide antisense RNA screening approach to identify genes important or essential for the growth of *Staphylococcus aureus* and found the *cld* homologue of this Gram-positive organism to be essential, too.

Expression and subcellular localization of LmCld

In about half of the collected fractions protein precipitated probably due to high concentration. Clear fractions containing LmCld were pooled and dialysed overnight against 20 mM HEPES (pH 7.4). Subsequently, protein solution was loaded on a HiLoad 26/60 Superdex 200 pg column (GE Healthcare) equilibrated with the same buffer. Fractions containing LmCld were pooled and concentrated to 1.3 mg/ml for crystallization trials.

Crystallization of LmCld

Initial crystallization trials were performed with different commercial and self-made screens at 22°C using the sitting-drop vapor diffusion technique and a nanodrop-dispensing robot (Phoenix™ RE).

Several conditions displayed hits, however most crystals were colorless and did not diffract. Nevertheless, some reddish crystals grew in 0.1 M Tris pH 8.5 and 2 M sodium formate (condition No. C5 of Hampton Research SaltRX Screen HT) at protein to mother liquor ratio 4:1. Prior to data collection, crystals were transferred into crystallization solution supplemented with 30% glycerol and were flash-frozen in a stream of nitrogen gas at 100 K using an Oxford Cryosystems Cryostream device.

Figure 2. **Flash frozen crystal of LmCld mounted in a cryo loop**.

Data collection and processing

Diffraction data were collected from a plate-shaped crystal with approximate dimensions of 200x100x40 µm³, at the ESRF beamline ID23-1 using 0.947 Å wavelength, to a maximum resolution of 1.8 Å. Diffraction data were integrated and scaled using XDS (Kabsch, 1993). Data collection statistics are summarized in Table 1.

Phasing, model building, refinement and validation of the structure

Phase problem was solved by molecular replacement using the Balbes webserver (Long *et al.*, 2008). The structure was refined with Refmac5, phenix.refine and TLSMD (Adams *et al.*; Murshudov *et al.*, 1997; Painter and Merritt, 2006). Manual model building was done in COOT (Emsley and Cowtan, 2004) while final validation of the models was performed with MOLPROBITY (Davis *et al.*, 2004). Refinement statistics are summarized in Table 1.

Structural analysis and superposition

Structure comparisons and superpositions were performed using the SSM server (Protein Structure Comparison service SSM at EBI), the program SUPERPOSE of the CCP4-package and Pymol (CCP4, 1994; DeLano, 2000; Krissinel and Henrick, 2004). The subunit interface of structures was analyzed using pdbsum (Laskowski, 2009) and the Protein Interfaces, Surfaces and Assemblies Service PISA at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) (Krissinel and Henrick, 2007).

Subunit interaction analysis was done with. Functional annotation was done with the Dali, MARKus server and ProFunc (Holm *et al.*, 2006; Laskowski *et al.*, 2005; Petrey *et al.*, 2009). All graphical representations were generated using PyMol (DeLano, 2000).

UV-vis spectroscopic analysis and heme content determination

UV-visible spectra were recorded on a Thermo Scientific Nanodrop 2000c Spectrophotometer at 22°C. Heme type and LmCld:heme *b* stoichiometry were determined by the pyridine hemochrome assay (Smith, 1975). The spectra were measured in 20% (v/v) pyridine and 0.1 M NaOH in 50 mM phosphate buffer (pH 7.0). Sodium dithionite was added before the measurement to a final concentration of 10 mM. The heme content was calculated using the molar extinction coefficient ε_{418}=191 500 M^{-1} cm^{-1} (Smith, 1975). For the reduced and oxidized spectra 10 mM sodium dithionite and 10 mM $K_3Fe(CN)_6$, respectively, were added.

Miscellaneous

Protein concentrations were determined by UV_{280nm} absorption. The extinction coefficient was calculated with ProtParam (Gasteiger *et al.*, 2005).

Results and Discussion

Structural analysis

LmCld crystallized as a pentamer in space group $P12_11$, with unit cell dimensions of 78.05 Å × 129.05 Å × 78.04 Å (Figure 3). The final model consists of 9717 protein atoms, 851 water molecules, has good

stereochemistry and R-factors (Table 1) The Wilson temperature factor is 19.2 Å^2. A clear electron density was observed for residues 5-110 and 123-252 of subunit A, for subunit B 5-111 and 123-253, for subunit C from 4-111 and 123-253, for subunit D from 5-111 and 123-251 and for subunit E from 6-112 and 123-252.

Table 1.
Data collection and refinement statistics for LmCld

A. Data Collection	
Beamline	ID23-1 (ESRF)
Wavelength (Å)	0.947
Resolution (Å)	44.8 – 1.9
	(1.95 - 1.90)
Space group	P12$_1$1
Unit cell (Å,°)	a = 78.05,
	b = 129.04,
	c = 78.04,
	β = 106.1
Molecules / a.u.	5
Unique reflections	216746 (15485)
Total reflections obs	414769 (29762)
Completeness (%)	93.9 (89.8)
R$_{meas}$^{b}	0.074 (0.365)
R$_{merge}$^{c}	7.1% (27.3%)
Multiplicity	1.9 (1.9)
I/σ(I)	8.01 (2.97)
B.Refinement	
R$_{cryst}$^{d}/ R$_{free}$^{e}	0.2387 / 0.3046
R.m.s.d. bonds (Å)	0.015
R.m.s.d. angles (º)	1.565

^{a}Values in parentheses are for the highest resolution shell.

$$^b R_{meas} = \sum_{hkl} \left[N/(N-1) \right]^{1/2} \sum_i \left| I_i(hkl) - \langle I(hkl) \rangle \right| \Big/ \sum_{hkl} \sum_i I_i(hkl)$$

$$^c R_{merge} = \sum_{hkl} \sum_i \left| I_i(hkl) - \langle I(hkl) \rangle \right| \Big/ \sum_{hkl} \sum_i I_i(hkl)$$

where $I_i(hkl)$ and $\langle I(hkl) \rangle$ are the ith and the mean measurements of the intensity of reflection hkl and N is the redundancy

^{d}R$_{cryst}$ = Σ |F$_o$-F$_c$| / Σ F$_o$

^{e}R$_{free}$ is the cross-validation R$_{factor}$ computed for the test set of reflections (5 %) which are omitted in the refinement process.
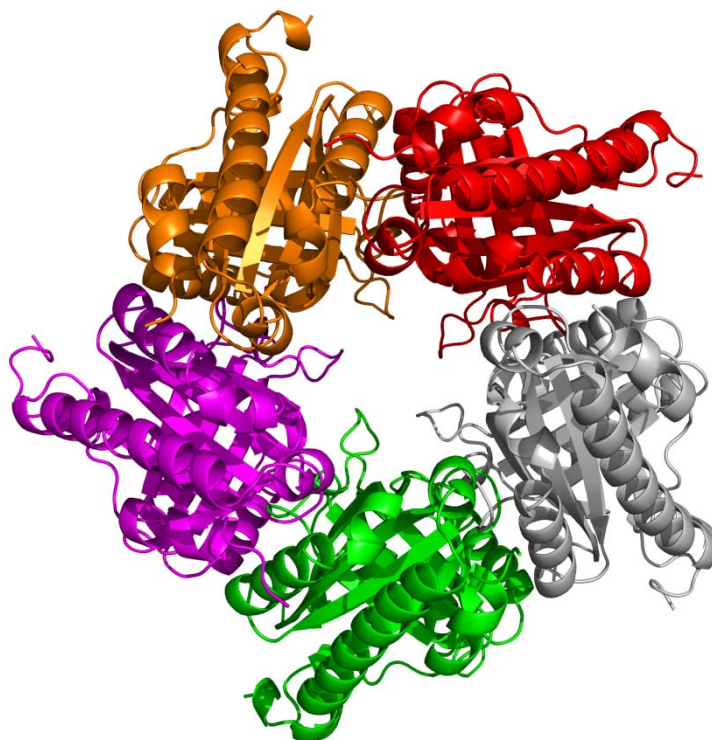
Figure 3. **Crystal structure of Cld from *Listeria monocytogenes***. Ribbon representation of the LmCld structure. Monomers are shown in a unique color. View on the N-terminal ferredoxin–like domain (top view).

Subunit structure

The subunit structure of LmCld is composed of two ferredoxin-like domains. In this common α+β protein fold two terminal β-strands hydrogen-bond to the central two β-strands, forming a four-stranded, antiparallel β-sheet covered on one side by two α-helices β4↑α1β1↓β3↑α2β2↓ (Figure 4A).

The same fold has also been observed in structures of Clds of *Thermoplasma acidophilum* (TaCld) (PDB: 3DTZ), *Geobacillus stearothermophilus* (GsCld) (PDB: 1VDH)*, Thermus thermophilus* HB8 (TtCld) (PDB: 1T0T), *Azospira oryzae* strain GR-1 (AoCld) (PDB: 2VXH) and *Ca. N. defluvii* (NdCld) (PDB:3NN1)) (Figure 4B).

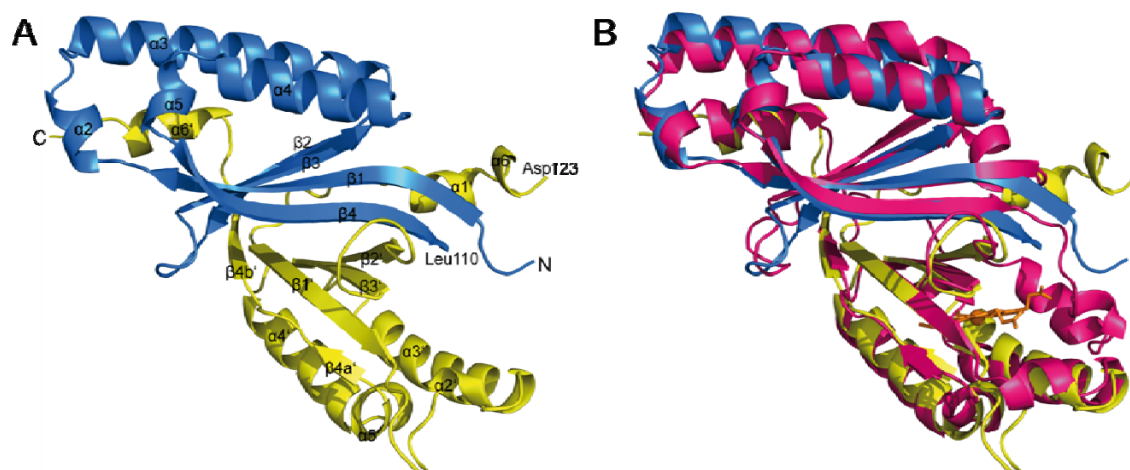Figure 4. **Comparison of subunit structures of LmCld and NdCld**. (A) The α helices and β strands are labeled according to the ferredoxin fold. The LmCld monomer consists of two ferredoxin-like domains, therefore secondary structure elements in the C-terminal domain follow the labeling of the N-terminal domain, with prime mark used as distinguisher. Additionally is the first ferredoxin like domain colored in marine the second domain in yellow. Parts of the structure between Leu110 and Asp123 could not be modeles , due to weak electron density in these region.(B) LmCld (same coloring as in A) superimposed with NdCld (shown in hotpink). The heme is shown in orange, with iron depicted as an orange sphere.

The two ferredoxin-like domains of LmCld superimpose with a root mean square deviation (r.m.s.d.) of 2.39 Å over 79 Cα atoms. Except of the last β-strand (β4'), which is divided into two parts by a loop, there are no major differences between N-terminal and C-terminal ferrodoxin-like domains of LmCld.

Comparison of Cld subunits

When compared to subunits of other Clds – Cld of *Nitrobacter winogradskyi* NwCld (Mlynek *et al.*, 2010); Cld of "*Candidatus* Nitrospira defluvii" (NdCld,PDB:3NN1); Cld of *Azospira oryzae* (AoCld, PDB: 2VXH); Cld of *Dechloromonas aromatica* (DaCld, PDB: 3M2Q and 3M2S); a Cld-like protein from *Geobacillus stearothermophilus* (GsCld, PDB: 1T0T); and to Cld-like protein from *Thermoplasma acidophylum* (TaCld, PDB: 3DTZ); and a Cld-like protein from *Thermus thermophilus* (TtCld, PDB: 1VDH), LmCld subunit superimposed with all other subunits with an r.m.s.d. of 1.63 Å over 178 pairs of structurally equivalent Cα atoms, revealing a high structural conservation of the subunit fold. LmCld showed the highest structural similarity to GsCld (r.m.s.d 0.339 Å) followed by TtCld, AoCld, TaCld and NdCld.

While in the crystal structures of validated Clds (NdCld, NwCld and AoCld), heme *b* was found bound in the C-terminal ferredoing-like domain, structures of GsCld, TaCld and TtCld were deposited in the PDB without heme cofactor. The main difference between the structures of validated Clds (NdCld, NwCld and AoCld) and Cld-like enzymes without bound heme (GsCld, TaCld and TtCld), is in the orientation of the loop region between β4 and α6 that is partly involved in binding of the prosthetic group. This loop also connects the two ferredoxin-like domains (Figure 4A, B). Besides this the putative heme binding cavity in Cld-like proteins is smaller. The third major difference validated Clds and Cld-like enzymes is the absence of helix α1 in Cld-like enzymes.

Quarternary structure

The LmCld pentamer forms a ring shaped assembly with an inner central pore of about 20 Å (Figure 3). Crystal structures of Cld-like proteins of *T. acidophilum*, *G. stearothermophilus, T. thermophilus* HB8 were also found in a pentameric architecture. The validated NdCld crystallized as a pentamer, while Cld from *A. oryzae* strain GR-1 (PDB: 2VXH) (AoCld) crystallized as a hexamer, but was shown to be a pentamer in solution (de Geus *et al.*, 2009). On the contrary, Cld from *N. winogradsky* (NwCld) is a functional dimer, displaying new quaternary structure found in some members of Cld-like family ((Mlynek *et al.*, 2010)).

In LmCld pentamer, the subunits are held together by a combination of hydrophobic and hydrogen-bonding interactions. About 35 residues of one subunit interact with 35 residues of the neighboring subunit making on average about 25 hydrogen bonding interactions and 215 non polar or hydrophobic contacts. No salt bridges have been found to stabilize the pentamer formation. The common interface between neighboring subunits is about 1650 $Å^2$, corresponding to about 24% of the total subunit solvent accessible area.

Biochemical characterisation for heme-binding

A detailed sequence analysis of Clds and Cld-like proteins suggested that in all enzymes of this superfamily, heme could be bound in the C-terminal ferredoxin domain (Ebihara *et al.*, 2005; Julius Kostan, 2010). LmCld was expressed in the presence of added hemin, similarly to the procedures used for purification of other Clds (de Geus *et al.*, 2008; Julius Kostan, 2010) for which high occupancy of the heme in the active site cavity was observed. However, in case of LmCld no electron density was observed for heme, despite the fact that crystals exhibit pale red color (Figure 2). Furthermore, several colorless crystals grew from initialization screens, indicating that heme is not covalently bound to protein and suggesting a weak interaction between the prosthetic group and the protein.

On the other hand heme analysis using the pyridine hemochrome assay showed peaks of dithionite-reduced LmCld at 418 nm (Soret peak), 525 nm (β max) and 557 nm (α max), suggesting that the heme incorporated is protoporphyrin IX (heme *b*) (Smith, 1975). Furthermore, the LmCld showed a Soret peak at 412 nm, which shifted to 426 nm upon addition of dithionite, corresponding to a reduced heme *b* moiety (Berry and Trumpower, 1987). The heme content was determined to be 0.17 heme *b* per LmCld monomer, while the Reinheitszahl value ($A_{Soret\,(412\,nm)}/A_{(280nm)}$) is 0.56.

To enrich protein on the heme *b* level, a reconstitution assay was performed. Purified Cld was mixed with 6 fold molar excess of hemin and incubated for 1 hour at different temperature (22°C, 35°C, 45°C, 55°C). Subsequently, the proteins were subjected to crystallization trials using the Hampton Research SaltRX screen where crystal were observed in the initial screening.

Interestingly, we did not obtain any crystal hits using reconstituted samples except for the sample prepared at 35°C. However, similarly to our previous observations, the crystals were again only pale red. A similar behavior was observed for the Cld of *Thermus thermophilus* (Ebihara *et al.*, 2005), where TtCld was found to contain no heme after over-expression in *E.coli*. In this case, reconstitution assay resulted in the heme *b* to protein ratio of 3:5 in solution. Unfortunately, the subsequent crystallization trials were unsuccessful.

Ebihara et al. suggest that a c-type heme with one thioether bond is formed upon binding to Cld. However structural and bioinformatics analysis shows no cysteine residue in a suitable position to form such a bond with the prosthetic group (Julius Kostan, 2010)

Structural analysis of low affinity heme binding pocket

Comparison of the structural models of the validated Clds and Cld-like proteins revealed remarkable differences in the heme-binding sites of these proteins.

Distal site

In a recent publication (Julius Kostan, 2010) Arg173 in NdCld was demonstrated to play a key role in (i) controlling of ligand and substrate access and binding and (ii) in chlorite dismutation reaction. The flexible side chain of Arg173 modulates the electrostatic potential and size of the active site entrance and might be involved in keeping transiently formed hypochlorite in place for final molecular oxygen and chloride formation. Furthermore, using a structure-based sequence alignment, it was shown that the position corresponding to Arg173 is conserved in all known active forms of Cld. Arg173 was therefore proposed as a molecular marker for Cld activity in yet uncharacterized Cld-like proteins. In LmCld and other Cld-like proteins Arg173 is replaced by glutamin (Gln187 in LmCld) or serine. (Figure 5).

```
Nitrobacter winogradskyi            48  ------------TPWQLRGVASHARYVERAE---KIALT----SVQAGLGRNEATRAALHPIRQSA----AWHEMTQDERRAIFEDKSHGIAASLKYLP-  123
Nitrobacter sp. Nb-311A             47  ------------TPWRLRGVTSHLRYVERAE---EVQAGLGRAEATRAALHPIRQSA----AWFELTQDERRKVFEDKSHGIAASLRFLP-  122
Bradyrhizobium japonicum            49  ------------NAWRLAGVFSSLRYTERAE---KQQLV---AVQAGLGRLEATSAALHPIRQSQ----AWFELTQERRRIFEDRSHGIASSLRFLP-  124
Limnobacter sp. MED105              48  ------------GTWMLQGFTSNVRYAERHE---INQLR---AKQEGLSRPASTCAALHPIKQNA----QWKALSQEERREIFEAKSHGTEIGLAYLP-  123
Pseudomonas aeruginosa              48  ------------GTWVLQGFTSNVRYAERHE---INQLR---AKQEELNRPTSSCAALHPIKHSP----AWHAMSQEEREIFEAKSHGTEIGLAYLP-  123
Gloeobacter violaceus               48  ------------AAWVLRGLTSNVRYAIRPE---VDALR---ERQPALARPEARCAALHPIRKTA----RWYELAQDERRDIFEETSHGTAIGMEFLP-  123
Cyanothece sp. PCC 7425             58  ------------SAWRLQGFASNIRYAIRTE---LEALQ---AVQPMLNRAEAILAVLHPIKQSA----AWQEMAQDERRDIFERESHGTAVGLEYLP-  133
Nitrosococcus mobilis               49  ------------AQWLLRGVTSNERYVTRSE---RAQLT---AKQPVLGRRQATCAAFHPIRQTA----SWGNLAQDERRMILEESSNHIKTGLKYLP-  124
Pseudomonas stutzeri                29  ------------TKWLLRGITSNERYVVREE---KDRLV---AKQFSLGRAEATCAALHPIRQNP----SWHGLSQDERRKIFEEQSRHIHIGLQYLP-  104
Klebsiella pneumoniae               49  ------------TKWLLRGITSNERYVVREE---KDRLV---AKQFSLGRAEATCAALHPIRQNP----SWHGLAQDERRKIFEEQSRHIHIGLQYLP-  124
Cupriavidus metallidurans           49  ------------AFWILRGITSNERYVMRAE---KNEIV---AKQQGLARPEATCGALHPIRQNA----AWFELTQDERRSVFE-QSRHVQIGLQYLP-  123
Ralstonia pickettii                 49  ------------AFWILRGITSNERYVMRAE---KNEIV---AKQQGLARPEATCGALHPIRQNA----AWFELTQDERRSVFE-QSRHVQIGLQYLP-  123
plasmid pAKD4                        1  .................MRAE---KNEIV---AKQQGLARPEATCGALHPIRQNA----AWFELTQDERRSVFE-QSRHVQIGLQYLP-   60
Janthinobacterium sp. Marseille     50  ------------WVLRGVTSNERVTKEE---KSRLI---AVQEGLGRTESTLAALHPIRQNA----SWHLLTQEERREILEERSHIQIGMAYLP-  123
Sorangium cellulosum                52  ------------TWVLGGVRSNERYVELEE---RRRLV---AVQEDLGRASSTQAALHPLRQSK----AWHALAQDERRAIFEARSRHIEIGLEYLP-  126

Ca. Nitrospira defluvii             85  AFMGTRLGRHLTSGGLLHGVSKKPTYVAGFP---ESMKT----ELQVNGESGSRPYAIVWPIKQDA----EWHALDQEARTALMQ---EHTQAALPYLK-  169
Azospira oryzae                     96  EFRSTTIGRNADVFETLVGVTKPLNYISKDK---SPGLNA--GLSSATYSGPAPRYVIVWPVKQNA----EWDNMSPEERLKEME---VHTTPTLAYLV-  182
Dechloromonas aromatica            127  EFRSTTVGRNADVFETLVGVTKPLNYISKDK---SPGLNA--GLSSATYSGPAPRYVIVWPVKQNA----EWDNMSPEERLKEME---VHTTPTLAYLV-  213
Dechlorosoma sp. KJ                127  EFRSTTVGRNADVFETLVGVTKPLNYISKDK---SPGLNA--GLSSATYSGPAPRYVIVWPVKQNA----EWDNMSPEERLKEME---VHTTPTLAYLV-  213
Dechloromarinus chlorophilus        45  EFRSTTVGRNADVFETLVGVTKPLNYISKDK---SPELNA--GLSSATYSGPAPRYVIVWPVKQNA----EWDNMSPEERLKEME---VHTTPTLAYLA-  131
Pseudomonas chloritidismutans       34  EFRSTTVGRNADVFETLVGVTKPLNYISKDK---SPELNA--GLTSATYSGPAPRYVIVWPVKQNA----EWDNMSPEERLKEME---VHTTPTLAYLV-  120
Pseudomonas sp. PK                  45  EFRSTTVGRNADVFETLVGVTKPLNYISKDK---SPELNA--GLSSATYSGPAPRYVIVWPVKQNA----EWDNMSPEERLKEME---VHTTPTLAYLV-  131
Dechloromonas agitata              120  DFRATRFGMNAEVTENLVGMTKDLNYITKDK---SPNLNA--GLTGATYRDATPRYAFVWPVKQNA----DWDNLTDEQRLKEME---THTLPTLANLV-  206
Dechloromonas sp. LT-1              45  DFRATRFGMYSDVTEQLLGMTKPLNYITKEK---SPDLNA--GLSSATYSAEAPPYAFVWPVKQSA----DWQNLTDAQRLKEME---THTLPTLGNLV-  131
Ideonella dechloratans             128  DFRATRFGMYSDVTESLVGITKALNYISKDK---SPDLNK--GLSGATYAGDAPRFAFMWPVKQNA----DWDNLTDEQRLKEME---THTLPTLPFLV-  214
Dechlorospirillum sp. WD            45  DWRATRLGMYSDVTENLVGITKALNYISKDK---SPELNA--GLSSATYSDSAPRYVIVWPVKQDA----AWDNMSDEQRLKEIE---VHTHGTLQYLV-  131
Magnetospirillum magnetobacticum   131  DFRATRFGMYSDVTENLVGITKALNYISKDK---SPDLNS--GLSSASYTGDAPRYAINWPVKQDA----AWDNRSDEERLKEME---VHTQPTLQYLV-  217

Geobacillus stearothermophilus      88  ALNKTKLADYLLPAYSYVSVVELSNYLA3GS---EDPYQIPEVRRRLYPILPKTNYICFYPMDDRRQGNDNWTMLSMEQRRELMR---AWGMTGRKYAG-  180
Listeria monocytogenes              88  RFNKLAIADYLLPTYSYISVVELSNYLASHMAGGDDPYQNKGVRARLYPALPPKKHICFYPMSPKRDGADNWYMLPMEERQQLIR---DHGLIGRSYAG-  183
Staphylococcus aureus               88  EFNKLRIADFLIPTYSYVSVIELSNYLAGKSD---EDPYENPHIKARLYPELPHSDYICFYPMNQRRNETYNWYMLTMEERQKLMY---DHGMIGRKYAG-  181

Thermus thermophilus                92  RLSRSAFARYLGRSYSFYSVVELGSQEKPLD---PESPY---VKPRLTPRVPKSGYVCFYPMNWRRQGQDNWTMLPAKERASLMK---AHGETGRKYQG-  181

Thermoplasma acidophilum            77  R-VQASMRPIAVSSFSSISIYDESPYNAMNK---KL---------EDSLRLPPLRYFVAYPMSDTP----DWYLLDFDTRKEIMH---EHIKMALNHPDE  156
Sulfolobus acidocaldarius           77  S-LISSGEGFLEEKLTLFSYFKPSPYIGGSAD---KL---------ASYLRLEPLRYFIAYPMKWSP----EWYLLPFEERKEIMD---EHIEIAKTHPDN  157


Nitrobacter winogradskyi           124  -AIARQLYHCRDIG---EPFDFLTWFEYAPEHATMFEDLVGVLRATEEWTY--VEREVDIRLAR-AI.........  183
Nitrobacter sp. Nb-311A            123  -AIARQLYHCRDIG---EPFDFLTWFEYAPEHATMFEDLVGVLRATEEWTY--VEREVDIRLAR-AV.........  182
Bradyrhizobium japonicum           125  -AIARQLYHCRDLG---GPFDFLTWFEFAPAHASLFEELVAMLRRTEEWTY--VEREVDVRVVK-EVLSA......  187
Limnobacter sp. MED105             124  -EIARQLHHSRDLG---EPFDFLTWFEFAPEHTDAFDELLVKLRTSEEWKY--VEREVDIRLVK-DSL........  184
Pseudomonas aeruginosa             124  -EIARQLHHSRDLG---EPFDFLTWFEYAPEHTAIFNKLLAQLRSSREWEY--VEREIDIRLVK-NV.........  183
Gloeobacter violaceus              124  -AVARRLHHCRDIG---EPFDFLTWFEYAPEHTHAFEDLLDRLRVTREWDF--VEREVDIRLEH-IDARI......  186
Cyanothece sp. PCC 7425            134  -GVARRLLHHCRDLG---EEFDFLTWFEPEHSSAFNLLLRMRASKEWEY--VEREVEVWLKR-L.........  192
Nitrosococcus mobilis              125  -AVARRLHHCRDLGGDAEPFDFLTWFEYAFSDSAAFDELVAELRASQEWTY--VDREIDMRLAR-DE........  187
Pseudomonas stutzeri               105  -AVARRLHHCRDLG-ENEPFDFLTWFSFPSDETAFNRLLAELRASVEWQY--VDREIDIRLVH-EPA........  168
Klebsiella pneumoniae              125  -AVARRLHHCRDLG-ESEPFDFLTWFSYFSFPSDEPGFNRLLAELRASVEWKY--VDREIDIRLVH-EPA........  187
Cupriavidus metallidurans          124  -AVARRLHHCRDLS-ENEPFDFLNWFEYAPIHEVEFNRLLSELRASEEWKY--VDREVDIRLTQ-AQV........  186
Ralstonia pickettii                124  -AVARRLHHCRDLS-ENEPFDFLNWFEYAPIHEVEFNRLLSELRASEEWKY--VDREVDIRLTQ-AQV........  186
plasmid pAKD4                       61  -AVARRLHHCRDLS-ENEPFDFLNWFEYAPIHEVEFNRLLSELRASEEWKY--VDREVDIRLTQ-AQV........  123
Janthinobacterium sp. Marseille    124  -AIARRLYHCRDIG-TPEPFDFLTWFSFPSDASQFDDLVAQLRTTHEWSY--IDHEIDIRLAR-RNG........  186
Sorangium cellulosum               127  -AVARRLYHSRDLG---GPFDFLTWFEFAEGDAGAFDELVGRLRETEEWSY--VEREIEVRLSR-A.........  185

Ca. Nitrospira defluvii            170  -TVKRRLYHST-GL---DDVDFITYFETE--RLEDFHNLVRALQQVKEFRHNRRFGHPTLLGTM-SPLDEILEKFAQ  238
Azospira oryzae                    183  -NVKRRLYHST-GL---DDTDFITYFETD--DLTAFNNIMLSLAQVKENKFHVRWGSPTTLGTI-HSPEDVIKALAD  251
Dechloromonas aromatica            214  -NVKRRLYHST-GL---DDTDFITYFETD--DLTAFNNIMLSLAQVKENKFHVRWGSPTTLGTI-HSPEDVIKALAD  282
Dechlorosoma sp. KJ                214  -NVKRRLYHST-GL---DDTDFITYFETD--DLTAFNNIMLSLAQVKENKFHVRWGSPTTLGTI-HSPEDVIKALAD  282
Dechloromarinus chlorophilus       132  -NVKRRLYHST-GL---DDT.........................  146
Pseudomonas chloritidismutans      121  -NV............................  122
Pseudomonas sp. PK                 132  -NVKRRLYHST-GL---DDT.........................  146
Dechloromonas agitata              207  -NVKRRLYHST-GL---DDTDFITYFETA--DLGAFNNIMLALAKVPENKYHVRWGSPTVLGTI-QSFDSVVNTLSM  275
Dechloromonas sp. LT-1             132  -NVKRRLYHST-GL---DDT.........................  146
Ideonella dechloratans             215  -NVKRRLYHST-GL---DDTDFITYFETN--DLGAFNNIMLSLAKVPENKYHVRWGNPTVLGTI-QPIENLVKTLSM  283
Dechlorospirillum sp. WD           132  -NVKRRLYHST-GL--ADA.........................  146
Magnetospirillum magnetobacticum   218  -NVKRRLYHST-GL---DDTDFITYFETN--DLGAFNNLLISLAKVPENKHHVRWGSPTVLGTI-QTFETVVKTLSM  286

Geobacillus stearothermophilus     181  -KVTQIITGSV-GL---DDFEWGVTLFSD--DALQFKKLVYEMRFDEVSAR-FGEFGSFFVGTRLPMENVSSFFHV.  248
Listeria monocytogenes             184  -KVQQIIGGSI-GF---DDYEWGVTLFSD--DALEFKRIVTEMRFDEASAR-YAEFGSFFIGNLLLSEQLSKLFTI.  250
Staphylococcus aureus              182  -RIKQFITGSV-GF---DDFEWGVTLFSD--DVLQFKKIVYEMRFDETTAR-YGEFGSFFVGHLINTNEFDQFFAIS  250

Thermus thermophilus               182  -EVMQVISGAQ-GL---DDWEWGVDLFSE--DFVQFKKIVYEMRFDEVSAR-YGEFGPFFVGKY-LDEEALRAFLGL  249

Thermoplasma acidophilum           157  KGIRSYTTYSF-GI---GDQEFVVLYEIP--DIAAWSRVTEKLREARARKW-IIKETPILLGRL-VDAGDIAGFLL.  224
Sulfolobus acidocaldarius          158  QGIRSYTTYSF-GI---ADYEFVVIYEAP--DLSKWINVVERLREAKARKW-VVSEEPILVGE----IGSLDIFLK.  222
```

Figure 5. **Sequence alignment of selected Cld homologs.**

Proximal site

In validated Clds a histidine residue from the α3' helix coordinates iron on the proximal side of the heme *b*. This histidine is hydrogen bonded to glutamic acid (Glu167 in NwCld), which in turn makes a hydrogen bond to lysine (Lys92 in NwCld)(Mlynek *et al.*, 2010). This lysine (Lys92 in NwCld) is also within hydrogen bonding distance to one propionate group of the heme *b*. This hydrogen bonding network increases the imidazolate character of the proximal histidine, thereby shifting the reduction potential of the heme iron to more negative values. This might be important in stabilization of higher heme oxidation state(s) involved in chlorite dismutation. Furthermore, on the proximal side of the heme *b*, at distance of 3.8 Å (in NwCld) from one of the propionate groups there is a pair of tryptophan residues (Trp96 and Trp97 in NwCld), which have their ring systems at right angle to each other. The first of these two residues has been suggested to play a role in the catalytic mechanism (Lee et al) (Figure 6) acting as the electron donor for the reduction of compound I to compound II.
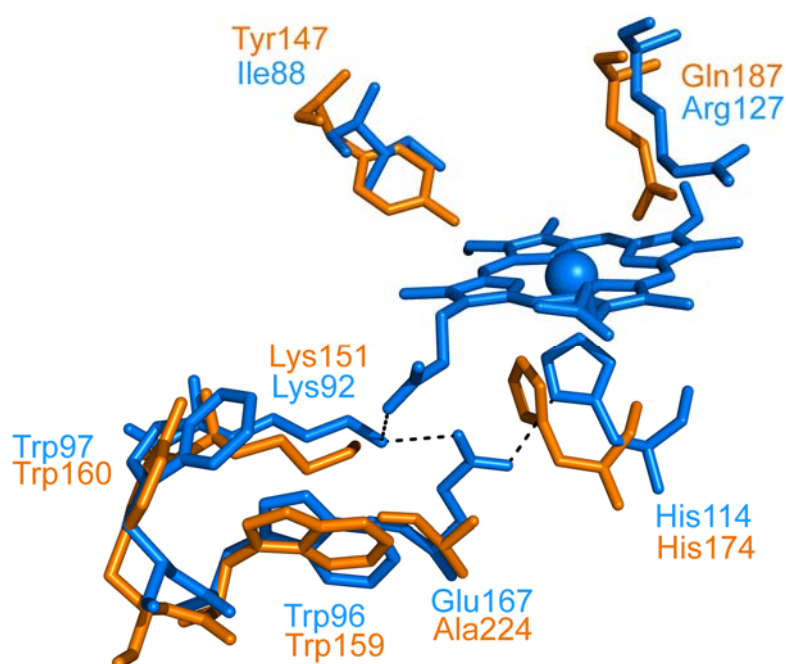
Figure 6. **Superpositioning of NwCld with LmCld.. NwCld is shown in marine and LmCld is shown in orange.** The figure shows selected residues in the active site of NwCld involved in heme *b* binding and in catalytic mechanism of chlorite dismutation. Two residues (Ile88 and Leu122) non conserved in Cld-like proteins, proposed to play a role in heme binding as discussed in the text, are also shown. The corresponding residues of NwCld are also shown. The heme iron is shown as blue sphere. Hydrogen bonding network in NwCld spanning from His114 to the propionate group of the heme is depicted with dashed lines.

In the structure of TtCld the imidazole ring of the proximal heme ligand (His172) is hydrogen-bonded (3.7 Å) to the side chain of Asp220. This interaction is similar to that observed in heme peroxidases which have a conserved proximal His-Asp pair. In LmCld the corresponding aspartic residue (Asp222) is in a different conformation, resulting in distance of 4.5 Å to the proximal histidine.

Additionally, the aforementioned Glu167, which stabilizes the proximal His114 and connects it *via* a network of hydrogen bonds to Lys92, is in LmCld replaced by an alanine without capacity to form a similar network (Figure 6). Trp97 which probably influences the ability of Trp96 to act as an electron donor, is replaced by a tyrosine. In the position corresponding to Ile88, LmCld and Cld-like proteins have a tyrosine residue (Figure 5). *In silico* mutation of Ile88 to tyrosine leads to a steric clash between the tyrosine side chain and the vinyl group of the heme moiety (Figure 6). Similarly, the second vinyl group in NwCld interacts with Leu122, while the Cld-like structures without heme have either an alanine or a glutamine in the corresponding position, leading to less favorable interactions with a putative heme *b* ligand. These subtle differences together with the different position of the loop between β4 and α1' which backbone interacts with heme propionate groups, suggest that if Cld-like proteins bind heme *b* they might bind it weaker and in a slightly different orientation than in the canonical Clds. Assuming these proteins still bind heme, it would be expected that their Cld activity is lower due to the differences in the key residues discussed above.

Analysis of electrostatics and solvent accessible surface

Similarly to all Clds and Cld-like proteins for which the structures are known, LmCld has a positive electrostatic potential on the surface around the entrance to the active site. For the validated Clds this is clearly important for attracting of the anionic substrate chlorite. In case of the Cld-like proteins this indicates that their substrate, if it is not chlorite, is also negatively charged. For the analysis of the solvent accessible surface we took as a reference structure for a validated Cld, - NdCld as it shows electron density over the all subunits, whereas in AoCld and NwCld one loop could not be modeled (REF). However results/conclusions should be applicable to all three Cld as the structures superpose very well (see section Comparison of Cld subunits) As a representative of Cld-like family we took TtCld because the loop connecting the two ferredoxin domains could be modeled in TtCld while the corresponding region was not observed in LmCld.

From Figure 7 it can be clearly seen that the heme is more buried in NdCld compared to TtCld. This is due to the fact that the loop (between XX and NN), connecting the two ferredoxin domains adopts a different orientation in LmCld and in TtCld compared to NdCld. This leads to an enlarged entrance to the putative active site.
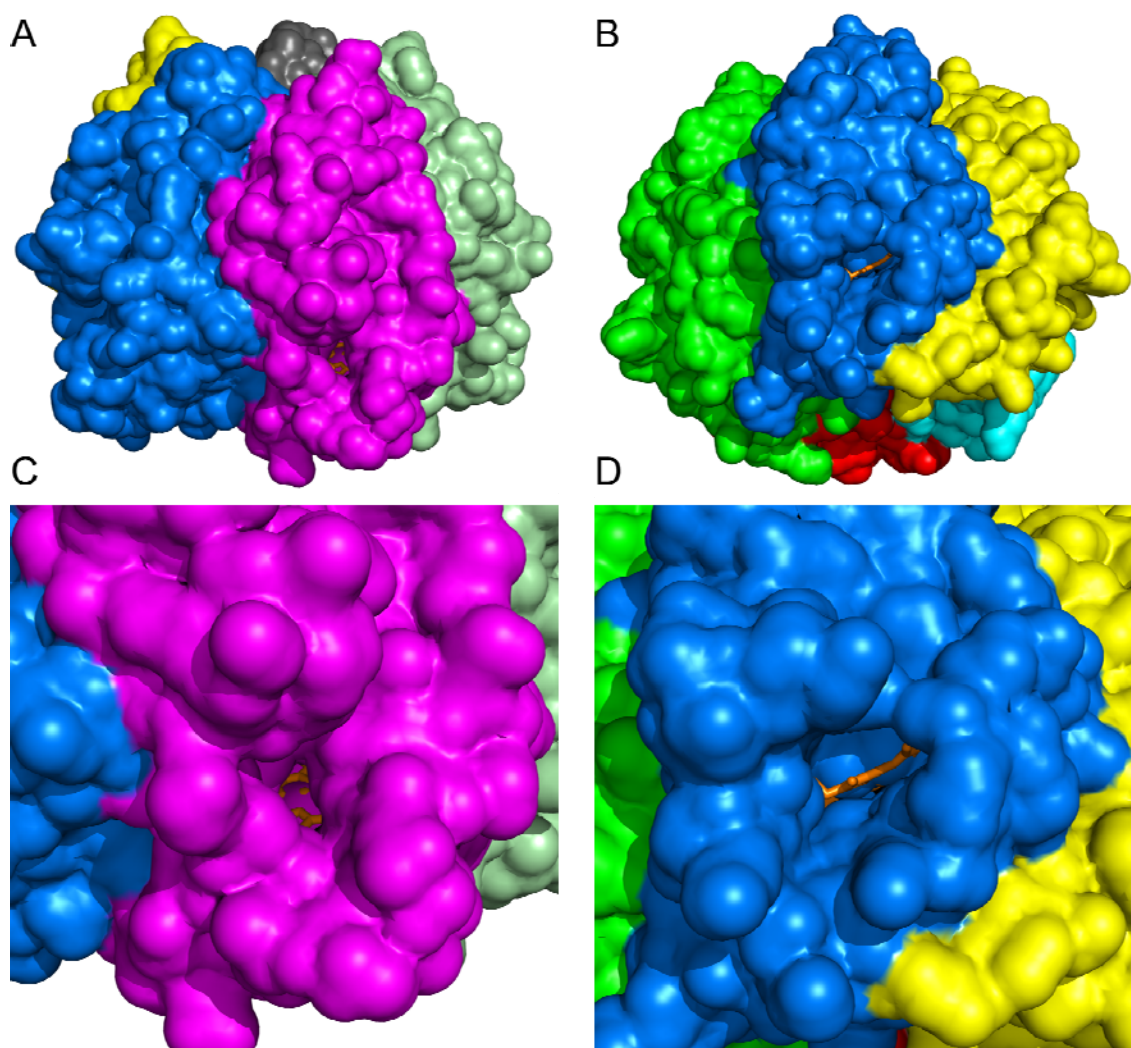
Figure 7. **Comparison of the solvent accessible surface of NdCld and TtCld and the substrate entrance.**
(A) The position and accessibility of the heme moiety in NdCld (overall view) is shown. .Monomers are shown in different colors. Heme is shown as orange sticks. (B) The position and accessibility of the heme moiety in TtCld (overall view) is shown. Monomers are shown in different colors. To build the structural of TtCld with heme in the putative active NdCld was superimposed onTtCld and heme coordinated read out.(C). Detailed view into the active site chamber through the putative substrate entrance and product exit channel NdCld (D). Detailed view into the active site chamber through the putative substrate entrance and product exit

Function of LmCld

Until now the function of LmCld and Cld-like enzymes remains unknown. LmCld was tested for chlorite dismutase-,halogenase-, peroxidase- and catalase- activity. Unfortunately the accurate protein/heme ratio wasn´t measured with this batch, but the protein solution was colored red and a soret band of 413nm was detected indicating that heme was bound. However just a weak catalase activity was seen for LmCld. Considering the abundance of a validated catalase in L. monocytogenes (Leblond-Francillard *et al.*, 1989) and the fact that the knockout of LmCld is lethal, LmCld is most likely not a catalase and fulfills a different function in *Listeria monocytogenes*.

It is common knowledge that the three-dimensional structure of a protein, which is tightly coupled to its molecular function, is more highly conserved than its amino acid sequence. Therefore structural comparisons may reveal biologically interesting similarities that are not detectable by comparing sequences. Besides sequence analysis we therefore performed structural comparisons with different programs.s

Structural comparison of LmCLd with Protein Data Bank using Dali

To compare structures in 3D we used the Dali server which compares the coordinates of a query protein structure against those in the Protein Data Bank (pdb).

The top hits (Z score) are the Cld´s 1T0T,= 1VDH, 3DTZ, 2VXH, a dye-decolorizing peroxidase (pdb entry 2GVK), a dye-decolorizing peroxidase (pdb entry 2HAG), Efeb - a component of the Efeuob bacterial iron transport system (pdb entry 2WX7), decolorizing peroxidase (DyP) (pdb entry 2D3Q), uncharacterized protein (pdb entry 2PGC), ORFan protein from *Chlorobium tepidum* with a ferrodoxin-like domain repeat (pdb entry 3GN6), putative monooxygenase (pdb entry 3F44), monooxygenase-like protein (pdb entry 3FJ2), self-compartmentalizing sulfur cycle metalloenzyme (pdb entry 2CB2). An extended table of comparison against a non-redundant subset of pdb entries (PDB90 is a representative subset of PDB chains, where no two chains are more than 90 % sequence identical to each other) is shown in table 3.

The structural similarities of the top hits to LmCld suggests that this proteins might have evolved from a common ancestral fold, known as the ferredoxin-like fold and can be tuned to fulfill different tasks. Interesting is that several of these structures (e.g. pdb entry 2IIZ; a peroxidase), also binds heme but in a slightly different orientation of the heme as seen in the validated Clds.

Table 3.

Result of DALI server analysis

No	Chain	Z	rmsd	lali	nres	%id P	Description
1	1t0t-V	35.4	0.6	234	243	59	MOLECULE: APC35880;
2	1vdh-A	33.1	1.3	234	247	47	MOLECULE: MUCONOLACTONE ISOMERASE-LIKE PROTEIN;
3	3dtz-D	22.6	2.3	213	223	15	MOLECULE: PUTATIVE CHLORITE DISMUTASE TA0507;
4	2vxh-B	20.1	2	195	230	16	MOLECULE: CHLORITE DISMUTASE;
5	2gvk-A	15	3	190	309	11	MOLECULE: HEME PEROXIDASE;
6	2hag-A	14.8	3	192	307	8	MOLECULE: MELANIN BIOSYNTHESIS PROTEIN TYRA, PUTATIVE;
7	2wx7-A	14.3	3.2	199	375	10	MOLECULE: PEROXIDASE YCDB;
8	2d3q-A	12.1	3.4	193	439	6	MOLECULE: DECOLORIZING PEROXIDASE;
9	2pgc-A	11.8	3.4	168	206	13	MOLECULE: UNCHARACTERIZED PROTEIN;
10	3gn6-C	10.2	3.6	178	301	9	MOLECULE: CT0912, ORFAN PROTEIN WITH A FERRODOXIN-LIKE
11	3f44-A	9.8	3.9	172	210	7	MOLECULE: PUTATIVE MONOOXYGENASE;
12	3fj2-A	9.5	3.3	148	170	9	MOLECULE: MONOOXYGENASE-LIKE PROTEIN;
13	2cb2-A	8.9	3.7	168	307	8	MOLECULE: SULFUR OXYGENASE REDUCTASE;
14	3bxv-A	8.5	3.8	170	308	8	MOLECULE: SULFUR OXYGENASE/REDUCTASE;

No: Rank of the match; Chain: PDB and chain identifier of the matched structure; Z: Z-score of the match. The list is ordered by decreasing Z-scores; rmsd: RMSD of the match; lali: Number of aligned positions; nres: Number of residues in matched structure; %id : Sequence identity of aligned positions; Description:Description of the matched structure (first COMPND line from PDB entry)

Structural comparison of LmCLd with Protein Data Bank using MARKus

For another structural comparison we used the MARKus server which compares the query protein structure against those in the Protein Data Bank (pdb).

The top hits are again the Cld´s 1T0T, 1VDH, 3DTZ, 2VXH, a decolorizing peroxidase (DyP) that catalyses the biological oxidation of anthraquinone derivatives (pdb entry 2D3Q), a dye-decolorizing peroxidase (pdb entry 2IIZ), Efeb - a component of the Efeuob bacterial iron transport system (pdb entry 2WX6), an uncharacterized protein (2PGC) (pentamer), ORFan protein from *Chlorobium tepidum* (pdb entry 3GN6), self-compartmentalizing sulfur cycle metalloenzyme (pdb entry 2CB2), putative monooxygenase (pdb entry 3F44), monooxygenase-like Protein (pdb entry 3FJ2). These structural homologs align on the N- and C terminal. Interestingly some hits just aligned with the C-terminal domain: a maize glutamine synthetase (pdb entry 2D3A), dissimilatory sulfite reductase (pdb entry 3C7B), yeast elongation factor 2 (pdb entry 1N0U).

Acknowledgment

References

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N *et al* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213-21.

Berry EA, Trumpower BL (1987). Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal Biochem* **161:** 1-15.

Camejo A, Buchrieser C, Couve E, Carvalho F, Reis O, Ferreira P *et al* (2009). In vivo transcriptional profiling of Listeria monocytogenes and mutagenesis identify new virulence factors involved in infection. *PLoS Pathog* **5:** e1000449.

CCP4 (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst* **D50:** 760-763.

Coates JD, Achenbach LA (2004). Microbial perchlorate reduction: rocket-fueled metabolism. *Nat Rev Microbiol* **2:** 569-80.

Davis IW, Murray LW, Richardson JS, Richardson DC (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **32:** W615-9.

de Geus DC, Thomassen EA, Hagedoorn PL, Pannu NS, van Duijn E, Abrahams JP (2009). Crystal structure of chlorite dismutase, a detoxifying enzyme producing molecular oxygen. *J Mol Biol* **387:** 192-206.

de Geus DC, Thomassen EA, van der Feltz CL, Abrahams JP (2008). Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of chlorite dismutase: a detoxifying enzyme producing molecular oxygen. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64:** 730-2.

DeLano WL (2000). *The PyMOL Molecular Graphics System on World Wide Web http://www.pymol.org*.

Drevets DA, Bronze MS (2008). Listeria monocytogenes: epidemiology, human disease, and mechanisms of brain invasion. *FEMS Immunol Med Microbiol* **53:** 151-65.

Ebihara A, Okamoto A, Kousumi Y, Yamamoto H, Masui R, Ueyama N *et al* (2005). Structure-based functional identification of a novel heme-binding protein from Thermus thermophilus HB8. *J Struct Funct Genomics* **6:** 21-32.

Emsley P, Cowtan K (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126-32.

Füreder (2009). Functional analyses of chlorite dismutase-like proteins from *Listeria monocytogenes* and *Nitrobacter winogradskyi PhD Thesis*.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD *et al* (2005). Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM (ed). *The Proteomics Protocols Handbook*. Humana Press. pp 571-607.

Glaser P, Frangeul L, Buchrieser C, Rusniok C, Amend A, Baquero F *et al* (2001). Comparative genomics of Listeria species. *Science* **294:** 849-52.

Goblirsch BR, Streit BR, DuBois JL, Wilmot CM (2009). Crystallization and preliminary X-ray diffraction of chlorite dismutase from Dechloromonas aromatica RCB. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65:** 818-21.

Hain T, Chatterjee SS, Ghai R, Kuenne CT, Billion A, Steinweg C *et al* (2007). Pathogenomics of Listeria spp. *Int J Med Microbiol* **297:** 541-57.

Holm L, Kaariainen S, Wilton C, Plewczynski D (2006). Using Dali for structural comparison of proteins. *Curr Protoc Bioinformatics* **Chapter 5:** Unit 5 5.

Ji Y, Zhang B, Van SF, Horn, Warren P, Woodnutt G *et al* (2001). Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. *Science* **293:** 2266-9.

Julius Kostan BS, Frank Maixner, Georg Mlynek, Paul Georg Furtmueller, Christian Obinger, Michael Wagner, Holger Daims, and Kristina Djinovic-Carugo (2010). Structural and functional analyses of the chlorite dismutase from the nitrite-oxidizing bacterium 'Candidatus Nitrospira defluvii': Identification of a molecular marker for enzyme activity. *JBC*.

Kabsch W (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* **26:** 795-800.

Krissinel E, Henrick K (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60:** 2256-68.

Krissinel E, Henrick K (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372:** 774-97.

Laskowski RA (2009). PDBsum new things. *Nucleic Acids Res* **37:** D355-9.

Laskowski RA, Watson JD, Thornton JM (2005). ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res* **33:** W89-93.

Leblond-Francillard M, Gaillard JL, Berche P (1989). Loss of catalase activity in Tn1545-induced mutants does not reduce growth of Listeria monocytogenes in vivo. *Infect Immun* **57:** 2569-73.

Long F, Vagin AA, Young P, Murshudov GN (2008). BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64:** 125-32.

Maixner F, Wagner M, Lucker S, Pelletier E, Schmitz-Esser S, Hace K *et al* (2008). Environmental genomics reveals a functional chlorite dismutase in the nitrite-oxidizing bacterium 'Candidatus Nitrospira defluvii'. *Environ Microbiol* **10:** 3043-56.

Marino M, Hoffmann T, Schmid R, Mobitz H, Jahn D (2000). Changes in protein synthesis during the adaptation of Bacillus subtilis to anaerobic growth conditions. *Microbiology* **146 (Pt 1):** 97-105.

Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C *et al* (1999). Food-related illness and death in the United States. *Emerg Infect Dis* **5:** 607-25.

Miroux B, Walker JE (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260:** 289-98.

Mlynek G, Sjöblom B, Kostan J, Füreder S, Maixner F, Furtmüller P *et al* (2010). Unexpected Structural and Phylogenetic Diversity of Chlorite Dismutases: A Catalytically Efficient Dimeric Enzyme from Nitrobacter winogradskyi
ISME J.

Murshudov GN, Vagin AA, Dodson EJ (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53:** 240-55.

Painter J, Merritt EA (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62:** 439-50.

Petrey D, Fischer M, Honig B (2009). Structural relationships among proteins with different global topologies and their implications for function annotation strategies. *Proc Natl Acad Sci U S A* **106:** 17377-82.

Schmid MW, Ng EY, Lampidis R, Emmerth M, Walcher M, Kreft J *et al* (2005). Evolutionary history of the genus Listeria and its virulence genes. *Syst Appl Microbiol* **28:** 1-18.

Smith KM (1975). *Porphyrins and metalloporphyrins*. Elsevier Scientific: Amsterdam.

Stenklo K, Thorell HD, Bergius H, Aasa R, Nilsson T (2001). Chlorite dismutase from Ideonella dechloratans. *J Biol Inorg Chem* **6:** 601-7.

Streit BR, DuBois JL (2008). Chemical and steady-state kinetic analyses of a heterologously expressed heme dependent chlorite dismutase. *Biochemistry* **47:** 5271-80.

Swalla BM, Cho EH, Gumport RI, Gardner JF (2003). The molecular basis of co-operative DNA binding between lambda integrase and excisionase. *Mol Microbiol* **50:** 89-99.

Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D *et al* (2009). The Listeria transcriptional landscape from saprophytism to virulence. *Nature* **459:** 950-6.

Chapter V

Summary/Zusammenfassung

Summary

In **Chapter II** we investigated the highly active Cld from Candidatus Nitrospira defluvii (NdCld), a key nitrifier in biological wastewater treatment, using a comprehensive structural, biochemical and bioinformatics approach. We cloned, heterologously expressed, crystallized, solved the phase- problem, determined the crystal structure and showed that functional NdCld is a homopentamer composed of two ferredoxin-like domains found in other Clds and Cld-like enzymes. To investigate the Cld function in more detail, site-directed mutagenesis of a catalytically important residue (Arg173) was performed and two enzyme mutants were structurally and biochemically characterized. Arginine 173 is demonstrated to play a key role in (i) controlling of ligand and substrate access and binding and (ii) in chlorite dismutation reaction.

The flexible residue modulates the electrostatic potential and size of the active site entrance and might be involved in keeping transiently formed hypochlorite in place for final molecular oxygen and chloride formation. Furthermore, using a structure-based sequence alignment, we show that the residue corresponding to Arg173 is conserved in all known active forms of Cld and propose it as a marker for Cld activity in yet uncharacterized Cld-like proteins.

In **Chapter III** by combining heterologous expression with structural and biochemical analyses, we identified the *N. winogradskyi* Cld protein, representing a new lineage in the superfamily of chlorite dismutases, as a novel and structurally distinct chlorite dismutase. This Cld is highly active, has just one ferredoxin-like domain and is a dimer. These results strengthened hypothesis we made in Chapter II on the reaction mechanisms, and suggest a yet unrecognized diversity of organisms potentially involved in the bioremediation of (per)chlorate and chlorite, and thus could open new perspectives for future research on these processes.

In **Chapter V** an investigation of the Cld-like protein of the well-described human pathogen *Listeria monocytogenes* was done. Our collaborators found that LmCld is essential for *L. monocytogenes* but shows no Cld-activity and no other function could be detected. That's why we set for a series of biochemical and structural studies with the aim to reveal the function of this enzyme in *Listeria monocytogenes*.

We heterologously expressed, purified, crystallized and determined the structure of LmCld. The LmCld crystallized as a pentamer and forms a ring shaped assembly. The subunit structure of LmCld is composed of two ferredoxin-like domains found also in other Cld-like enzymes. A comparison of the structural models of the validated Clds and Cld-like proteins revealed remarkable differences in the heme-binding sites of these proteins. The proximal part is missing a hydrogen bonding network which could shift the reduction potential of the heme iron to more negative values which is important in stabilization of higher heme oxidation state(s) involved in chlorite dismutation. Additionally glutamine replaces argenine at the distal site. We could therefore strengthen the hypothesis that ARG is a

molecular marker for Cld activity in yet uncharacterized Cld proteins. Furthermore we found that if Cld-like proteins bind heme *b* they might bind it weaker and in a slightly different orientation than in the canonical Clds. A structural comparison suggests that this proteins might have evolved from a common ancestral fold, known as the ferredoxin-like fold and this fold can be tuned to fulfill different tasks. So far we couldn´t reveal the function of LmCld. However we keep working on this project as it is a target for structure based drug design, because it was shown to be essential for *Listeria monocytogenes*.

Zusammenfassung

Chlorite-dismutasen (Cld) sind Enzyme die toxisches Chlorite zu Chlor und Sauerstoff abbauen können. Der genaue Reaktionsmechanismus ist zurzeit aber noch unklar. Sie wurden 1996 entdeckt und bis vor kurzem glaubte man, dass diese nur (Per)chlorate-reduzierenden Bakterien vorkommen. Umso erstaunlicher war es, dass Cld 2008 in vielen Stämmen von Archaeen und Bakterien (einige davon sind human-pathogen wie zum Beispiel *Bacillus anthracis, Listeria monocytogenes*) gefunden wurden. Somit ergeben sich drei Hauptgründe Chlorite dismutasen zu untersuchen.

1. Der Reaktionsmechanismus ist hoch interessant – jedoch welche Aminosäuren in der Katalyse involviert sind bedarf noch detaillierter Untersuchungen.

2. Per(chlorat) und Chlorit werden - obwohl toxisch noch immer in vielen Bereichen z.B. als Bleichmittel, Raketentreibstoffe und Desinfektionsmittel eingesetzt. Biologischer Abbau dieser Substanzen ist die effizienteste Methode diese Umweltschadstoffe zu entfernen.

3. Da kein Cld homologes Gen im Menschen vorhanden ist und Cld in vielen human pathogenen Bakterien essentiell sind, stellen Cld ein neues Ziel für Struktur-basierte Medikamentenforschung dar.

Im zweiten Kapitel untersuchten wir die hoch aktive Chlorite-dismutase von Candidatus Nitrospira defluii, ein Schlüssel-Bakterium in Kläranlagen und dem globalen Stickstoffzyklus. Wir zeigten, dass das Enzyme aus 5 Untereinheiten aufgebaut ist und jede Untereinheit aus zwei Ferredoxin-ähnlichen Domainen besteht. Um den Reaktionsmechanismus genauer zu untersuchen mutierten wir die Aminosäure Arginine173 und konnten nachweisen dass Arginine173 (i) eine Schlüsselrolle in der Reaktion (ii) wie auch bei der Bindung von Chlorit an das aktive Zentrum hat. Mit diesem Ergebnisse ist es nun möglich anhand der Aminosäuren-Sequenz festzustellen, ob es sich um eine aktive Chlorite-dismutase handelt oder um ein Chlorite-dismutase ähnliches Enzyme. Bis jetzt weiß man noch nichts

über die Rolle von Chlorite-dismutasen ähnlichen Enzymen. In Kapitel IV wird genauer darauf eingegangen.

Im dritten Kapitel konnten wir die Hypothesen, welche wir im zweiten Kapitel über den Reaktionsmechanismus und Häm-Protein-Interaktionen aufstellten, anhand der hoch aktiven Cld von Nitrobacter winogradsky, einem zentralem Bakterium in Kläranlagen und dem Stickstoffzyklus, stärken. Zu unserer großen Überraschung stellte sich heraus, dass diese Cld eine bis jetzt nicht in der Proteinfamilie gesehene Quartärstruktur hat. Unsere Ergebnisse führen zu der Annahme, dass viel mehr Mikroorganismen als bisher gedacht (Per)chlorat und Chlorit abbauen können, was neue Perspektiven im Schadstoffmanagment öffnet.

Im letzten Kapitel analysierten wir das Chlorite-dismutasen ähnliche Enzym von dem human pathogenen Bakterium *Listeria monocytogenes*. Unsere Kooperationspartner konnten zeigen, dass die Chlorite-dismutase essential für *Listeria m*onocytogenes ist. Die Rolle des Cld-ähnlichen Enzymes konnte jedoch nicht gefunden werde. Nachdem die Tertiärstruktur von Proteinen stärker konserviert ist als die Primärstruktur hofften wir mit Hilfe der 3-D-Struktur die Funktion der Cld in *Listeria monocytogenes* entschlüsseln zu können. Die Quartärstruktur besteht wie bei NdCld aus 5 Untereinheiten, wobei wiederum jede Untereinheit aus 2 Ferredoxin ähnlichen Domänen aufgebaut ist. In der vermeintlichen Hämbindungsstelle sahen wir jedoch viele Unterschiede im Vergleich zu den hoch aktiven Cld. Auch ist das katalytisch wichtige Argenine in Cld-ähnlichen Proteien meist durch ein Glutamin ersetzt und auf der proximalen Seite fehlt ein Netzwerk aus Wasserstoffbrückenbindungen. Wir vermuten deshalb, dass Cld-ähnliche Enzyme Häm nicht binden können oder nur in einer anderen Orientierung als in hoch aktiven Cld. Ein Vergleich mit anderen Strukturen lässt uns zu dem Entschluss kommen, dass die Ferredoxin-ähnliche Domäne sehr früh in der Evolution entstanden ist und „getuned" werden kann um verschiedene Reaktion zu machen. Das große Ziel - die Funktion von der Listerien Cld zu entschlüsseln haben wir bis zum jetzigen Zeitpunkt nicht erreicht. Wir arbeiten aber weiterhin an diesem hochinteressanten Projekt und sind zuversichtlich die Funktion von Cld-ähnlichen Enzymen zu finden.

Insgesamt erweitern alle drei Kapitel dieser Diplomarbeit unser bisheriges Wissen über Chlorite-dismutasen in vielen Bereichen und eröffnen neue Perspektiven für zukünftige Forschungen in diesen Gebieten.

Appendix

Danksagung

Auf allen meinen bisherigen Wegen hatte ich immer eine Familie bei oder hinter mir welche mich unterstützte und mir die Freiheit gab meine Wege zugehen. Aus vollem Herzen danke ich Euch dafür.

Obwohl ich oft an meinen Zielen gescheitert bin – das wichtigste Ziel – glücklich zu sein habe ich erreicht. Einen großen Beitrag zu meinem erfüllten Leben leisteten die vielen Menschen die ich auf meinen Wegen traf. Ob im Sport, im Studium oder beruflich, bei allen diesen Personen möchte ich mich auch herzlich bedanken.

Curriculum Vitae

Personal

Name	Mlynek Georg
Date of Birth	October 1, 1980
Place of Birth	Linz, Austria
Citizenship	Austrian

University

Oct. 2003-Dec. 2010	Diploma studies of **Molecular Biology** at the University of Vienna
Oct. 2006-Mar. 2008	Bachelor studies in Business and Economics at the WU in Vienna
Mar.2005-Oct.2008	Bachelor studies in oenology at the University of Applied Sciences in Vienna
March 2001-July 2010	Diploma studies of **Microbiology / Genetics** at the University of Vienna

Military Service

Oct. 2000 – Sep. 2001	Heeresleistungsportzentrum at the Upper-Austrian Olympic base in Linz

Grammar school

Oct.1995 – June 2000	Bundesoberstufenrealgymnasium für Leistungssportler in Linz
Oct.1991 – July 1995	BRG Hammerling in Linz